THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En mécanismes des interactions parasitaires pathogènes et symbiotiques

École doctorale [GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau]
École doctorale [EDEG – Économie et Gestion]
Portée par

Unité Mixte de Recherche Biologie et Génétique des Interactions Plante-Parasite et l'unité de mycologie du laboratoire de la santé des végétaux de l'ANSES

Diagnostic et inférence de l'histoire évolutive des lignées endémiques et pandémiques de *Pyricularia oryzae* causant la pyriculariose du riz, du blé, et d'autres Poacées sauvages

Présentée par Maud THIERRY
Le 03 décembre 2019

Sous la direction de Didier THARREAU et Renaud IOOS

Devant le jury composé de

Mme Claire NEEMA, directrice de recherche, Supagro - UMR BGPI	**Présidente du jury**
Mme Marie-Agnès JACQUES, directrice de recherche, INRA Angers - UMR IRHS	**Rapporteur**
M. Pascal FREY, directeur de recherche, INRA Nancy - UMR IAM	**Rapporteur**
M. Philippe REIGNAULT, professeur des universités, ANSES Angers - LSV	**Examinateur**
M. Thomas BALDWIN, chercheur, BioGEVES - pôle détection	**Examinateur**
M. Didier THARREAU, directeur de recherche, CIRAD, UMR BGPI	**Directeur de thèse**
M. Renaud IOOS, directeur de recherche, ANSES Nancy, Unité de mycologie	**Co-directeur de thèse**
M. Pierre GLADIEUX, chargé de recherche, INRA Montpellier - UMR BGPI	**Membre invité (co-encadrant)**
Mme Elisabeth FOURNIER, directrice de recherche, INRA Montpellier - UMR BGPI	**Membre invité (co-encadrant)**

UNIVERSITÉ DE MONTPELLIER — Montpellier SupAgro — cirad — anses agence nationale de sécurité sanitaire alimentation, environnement, travail

Remerciement

Je remercie tout d'abord mes quatre encadrants Didier THARREAU, Renaud IOOS, Pierre GLADIEUX et Elisabeth FOURNIER pour m'avoir donné la chance de m'investir dans ce sujet de thèse qui m'a beaucoup appris, sur *Magnaporthe* bien sûr, mais aussi sur une multitude d'autres domaines inconnus jusqu'alors. Ça a été une expérience très enrichissante. Merci pour toutes vos corrections, votre disponibilité et vos conseils.

Je remercie Marie-Agnès JACQUES et Pascal FREY pour avoir accepté d'être les rapporteurs de cette thèse, ainsi que, Thomas BALDWIN, Philippe REIGNAULT et Claire NEEMA pour avoir accepté de faire partie du jury.

Je remercie également les membres de mon comité de thèse, Cyril DUTECH, Franck PRUGNOLLE, Marc-Henri LEBRUN et Stéphane DE MITA pour leur écoute et leurs suggestions.

Merci ensuite à ceux qui dans mon équipe Montpellieraine m'ont apporté leur aide au quotidien et leurs sourires. Merci à Joëlle pour ton aide plus que précieuse au laboratoire, pour tous tes conseils et tous les stocks STC, repiquages, lectures de symptômes et j'en passe. Merci à Henri pour m'avoir aidé à découvrir le monde des serres de BGPI, pour tous les semis et arrosages, pour m'avoir fait découvrir la Camargue et également pour ton aide concernant le « mystère des riz jaunes ». Merci à Sandrine pour tes coups de main au labo, tes conseils et tes encouragements tout au long de cette thèse. Merci à Sonia, pour ton aide au laboratoire même lors de mesures de croissance ou de panne des chambres de culture le week-end. Merci à Geneviève pour avoir toujours su m'aider avec le coté administratif dans lequel j'aurai été perdu sinon. Merci à Sébastien pour les innombrables fois ou tu as résolu mes problèmes informatiques (satané linux !) et pour toutes les analyses que tu m'as aidé à faire. Et enfin merci à François pour tes conseils en statistiques et les moments conviviaux que tu n'as pas hésité à partager avec les jeunes.

Merci à Loïc, Remy, Michel, Katia, Charlotte, Sonia et Joëlle pour leur aide en ce qui concerne les serres, les chambre de culture, la laverie et les autoclaves.

Merci à tous ceux qui a BGPI m'ont apporté leur aide ponctuelle et leur bonne humeur que ce soit pour emprunter du milieu de culture à la dernière minute ou quelques grains de blé, partir à la recherche de kit d'extraction, aller courir entre midi et deux, changer un pneu crevé devant le bâtiment, recharger

ma carte désespérément vide à la cantine ou discuter sous les hottes, les microscopes ou à la machine à café.

Merci à toute l'équipe de Nancy pour son accueil pendant les quelques mois que j'ai passé là-bas. En particulier merci à Renaud, encore, pour ta convivialité, ta grande disponibilité et ton aide pendant tout mon séjour nancéen. Merci à Jaime d'avoir été mon coloc de bureau pendant quatre mois. Merci à Sylvie pour ton amitié et ton aide pour tout le coté administratif. Merci à Violaine et Axel pour vous être investi dans certaines parties de ce projet de thèse. Et enfin merci à Céline, Cécile, Jacqueline, Véronique, Isabelle, Nathalie et Carole pour votre gentillesse et pour tout ce que vous avez pu m'apprendre du laboratoire.

Merci à tous les thésard, stagiaires, postdocs de BGPI et tous leurs +1, pour leur amitié et pour toute la bonne humeur qu'ils ont apporté pendant cette aventure. Merci à Emeline pour avoir brisé la solitude de mon bureau en y apportant ta joie, ton humour et ton amitié pendant ces trois années. Merci à Maëlle pour toutes les discussions, courses, séances de natation intensive, et discussions (encore !) qu'on a pu avoir pendant trois ans et qui continueront encore longtemps je l'espère. Merci à Coralie pour ta bonne humeur permanente et ta positive-attitude qui montre l'exemple. Merci à Edwige pour avoir été à l'origine de ce groupe de jeunes qui a tellement compté. Merci à Faustine pour tes points de vue toujours intéressants dans toute sorte de conversation. Merci à Valentin pour tes poèmes, charades, blagues, scotch, chocolats et j'en passe qui apportent leur dose de rire et de surprise chaque jour dans le bureau 124. Merci à Florian d'avoir été aussi un membre fort apprécié du bureau 124 et de m'avoir sorti de beaucoup de galères linuxiennes. Merci à Thomas (x2), Mathias et Remi pour les micro-pauses café dix fois par jour toujours appréciées. Merci à Jérémy, Jérémy, Guillaume, Albin, Kevin, Pascal, Anne, Marie, Mireille, Adeline et Audrey pour toutes les discussions, repas, pauses café, apéros et moments de rires passés ensemble au cours de ces trois années.

Merci à tous ceux qui a BIOGER ont participé à lancer ma carrière scientifique. Marc-Henri et Lilian qui m'ont accueilli en stage et m'ont fait découvrir pour la première fois un laboratoire de recherche.
Anaïs, Johann, Clémentine et toute la cellule du club G.R.O.S. avec qui je pense avoir passé 5 mois à rigoler en permanence et avec qui on rigolera encore longtemps.

Merci à toutes les amies qui me sont chères et qui sont là pour moi depuis tellement d'années Lauriane, Oriane, Cyrielle, Emilie, Manue, Léa, Camille, Sandy, Nicole et Coralie et qui n'hésitent pas régulièrement à me redemander sur quoi je travaille déjà (le blé c'est ça ?).

Merci à toute ma famille qui est certainement un peu à l'origine de cette vocation dans la recherche. Merci surtout à ma maman qui m'a toujours poussé dans mes études et soutenus à toutes leurs étapes. Merci aussi à Gilles et Marie avec qui ça a été un grand plaisir de passer du temps ces dernières années.

Et enfin merci à Yohann qui a partagé toutes les étapes de cette thèse, pour nos conversations (parfois scientifiques), pour tes conseils toujours avisés et pour toujours me soutenir dans mes projets aussi nombreux soient-ils. Cette thèse n'aurait vraiment pas été la même sans toi !

Table des matières

Liste des figures

Liste des tableaux

INTRODUCTION

Les champignons phytopathogènes sont des menaces pour les cultures agricoles et pour l'environnement. Connaitre au mieux ces microorganismes est essentiel pour pouvoir mettre en place des méthodes de lutte adaptées. Un certain nombre de questions sont récurrentes dans l'étude des agents pathogènes.

- Quels sont les symptômes de la maladie engendrée par cet agent pathogène ? Quel est son impact ? Quel est son cycle épidémique ?
- Où la maladie est-elle actuellement déclarée ? Où elle est apparue ? Comment s'est-elle propagée ? Peut-on prédire où elle peut se propager par la suite ?
- Comment évolue l'agent pathogène ? Peut-on prédire comment il peut évoluer ? Sera-t-il capable de contourner de nouvelles résistances variétales mises sur le marché ou devenir résistant à de nouveaux pesticides utilisés ?
- L'espèce est-elle homogène ou peut-on décrire en son sein des groupes d'individus présentant des caractéristiques particulières qu'il faudrait alors traiter de façon différente ?

Dans le cadre de cette thèse l'étude de la structure des populations d'une espèce fongique phytopathogène modèle, *Pyricularia oryzae*, permet d'éclairer certaines de ces questions, d'améliorer les méthodes de lutte contre ce parasite et enfin permet d'approfondir les connaissances générales sur l'évolution et la propagation des champignons phytopathogènes.

1- Structure de la biodiversité des agents pathogènes fongiques : des populations différenciées aux radiations adaptatives

La structure génétique des populations est définie comme l'organisation des variations génétiques au sein des populations et entre les populations. Cette structure est donc décrite grâce à l'étude de marqueurs génétiques, de leurs fréquences et de leurs associations dans les génomes des individus et dans les populations d'une même espèce. Comme le décrit M. Slatkin en 1987 les variations phénotypiques et les différences de fréquences génétiques entre populations au sein d'une même espèce résultent d'un équilibre entre des forces entrainant une différenciation génétique des populations et au contraire des forces permettant leur homogénéisation. Ici nous proposons de décrire certains des processus qui façonnent la structure génétique des populations, c'est-à-dire, qui permettent de comprendre comment se mettent en place et se maintiennent les barrières aux flux de gènes conduisant à la subdivision génétique d'une population initialement indifférenciée.

1.1- Recombinaisons : brassage génétique

Les phénomènes de recombinaison sont à la base de l'homogénéisation des populations. Ceux-ci permettent un brassage génétique entrainant l'association de marqueurs génétiques (un allèle donné à un locus donné) apparus indépendamment (chez des individus différents) au cours de l'évolution ou au contraire en permettant la dissociation de marqueurs initialement liés génétiquement. Dans les espèces fongiques ces évènements de recombinaison peuvent avoir lieu lors de différents processus biologiques.

Le brassage de l'information génétique à principalement lieu lors de la reproduction sexuée. Lors de celle-ci ont lieu des recombinaisons entre les génomes parentaux entrainant l'apparition dans leur descendance de nouvelles combinaisons alléliques. Les recombinaisons ont lieu lors d'événements de méiose, par appariement entre chromosomes homologues suivis d'échanges de matériel entre points de recombinaison (*crossing over*). Ces événements de recombinaison peuvent être détectés par l'analyse des déséquilibres de liaison, association préférentielle entre allèles à différents loci. Chaque individu issu de la reproduction sexuée porte donc aux différents loci le long du génome des allèles provenant de l'un ou l'autre de ses parents, et ces associations alléliques changent d'un événement de méiose à un autre. La reproduction sexuée est apparue tôt dans l'évolution des eucaryotes et a été conservée dans la majorité des espèces actuelles (Hofstatter and Lahr, 2019). Chez les espèces fongiques, il existe une grande diversité de cycles de reproduction sexuée dans la nature mais certaines caractéristiques générales peuvent tout de même être notées. Les champignons peuvent être isogames ou anisogames. L'isogamie signifiant que tous les gamètes possèdent une morphologie

identique ne permettant pas une distinction entre gamète mâle et gamète femelle et l'anisogamie signifiant que les gamètes ont des morphologies différentes (des différences de tailles). La plupart des champignons ascomycètes filamenteux sont anisogames tandis que les levures et les champignons basidiomycètes sont généralement isogames (Billiard et al., 2011; Debuchy et al., 2010). De plus, parmi les espèces anisogames, tous les individus sont capables de produire les deux types de gamètes. Ceci signifie qu'il n'existe pas d'individus de différents sexes au sein des champignons (Billiard et al., 2011; Coelho et al., 2017). Toutefois, la plupart des espèces fongiques sont hétérothalliques, impliquant qu'une syngamie (fusion des gamètes) ne peut avoir lieu qu'entre des gamètes de types sexuels opposés (Billiard et al., 2011; James, 2015; Nieuwenhuis et al., 2013). Le type sexuel peut alors être gouverné par un seul locus (systèmes bipolaires) ou par deux loci non liés génétiquement (systèmes tétrapolaires) (Bennett and Turgeon, 2017). Dans le cas d'un système tétrapolaire, les deux loci doivent être de types sexuels opposés pour que la reproduction sexuée puisse avoir lieu (Casselton and Olesnicky, 1998; Lee et al., 2010). Les loci gouvernant les types sexuels encodent des phéromones génétiquement liées à des récepteurs de phéromones ou des facteurs de transcription de type homéodomaine qui permettent de déterminer la viabilité de la descendance (Coelho et al., 2017).

Un second processus, peut-être plus anecdotique, est la parasexualité, qui peut exister chez certaines espèces fongiques. Ce processus a lieu lorsque les hyphes de deux individus dits « végétativement compatibles » fusionnent pour former une anastomose. Toutefois, la compatibilité végétative est régie par de nombreux loci (onze loci chez *Neurospora crassa*, neuf chez *Podospora anserina*, huit chez *Aspergillus nidulans*, et au moins cinq chez *Cryphonectria parasitica* – Anagnostakis, 1982; Bernet, 1965; Jinks and Grindle, 1963; Perkins, 1988) au niveau desquels les allèles doivent tous être identiques chez les deux individus pour qu'une fusion des hyphes puisse avoir lieu. Seuls les individus génétiquement très proches seront donc végétativement compatibles et la très grande majorité des paires d'individus seront végétativement incompatibles au sein d'une population. Dans cette hyphe fusionnée, les matériels nucléaires et cytoplasmiques des deux individus sont mis en commun formant des hétérocaryons, cellules multinucléées dont les noyaux sont génétiquement différents (Glass, 2004). Des phénomènes de recombinaison au sein de ces hétérocaryons ont pu être démontrés chez plusieurs espèces fongiques suite à la fusion de noyaux cellulaires (*Aspergilus nidulans*, *Neurospora crassa*, *Candida albicans* ou certains champignons mycorhiziens arbusculaires – Chen et al. 2018; Orr-Weaver et Szostak 1985; McManus et Coleman 2014). Cette parasexualité a été démontrée en laboratoire mais son existence dans des conditions naturelles ne l'est pas (Nieuwenhuis and James, 2016).

Un troisième processus, beaucoup mieux documenté que la parasexualité chez les champignons, est le transfert horizontal de gène, défini comme le transfert de matériel génétique entre individus

autrement que par hybridation (c'est-à-dire autrement que par la reproduction sexuée ou parasexualité). Les transferts horizontaux sont des phénomènes fréquents chez les bactéries. Chez les eucaryotes, et chez les champignons en particulier, ce phénomène a également été décrit chez plusieurs espèces. Des transferts horizontaux entre le règne fongique et d'autres règnes du vivant ont été mis en évidence dans plusieurs études (Taylor et al., 2017) : entre plante et champignon (Richards et al., 2009) ; entre champignon et oomycète (Richards et al., 2011) ; champignon et bactérie (Hahn et al., 2005) ; champignon et moustique hôte (Wang et al., 2016). Des transferts horizontaux entre espèces fongiques sont aussi possibles. Un des exemples marquants est le transfert du gène *ToxA*, codant pour une toxine, entre différentes espèces de champignons ascomycètes. Ce gène confère une virulence immédiate vis-à-vis des variétés de blé portant le gène se sensibilité à cette toxine, le gène *Tns1* (Friesen et al., 2006; Taylor et al., 2017). Autre exemple, le transfert horizontal d'un grand complexe de gènes produisant des métabolites secondaires parmi lesquels des antibiotiques tels que des pénicillines (Kroken et al., 2003; Taylor et al., 2017).

Dans les populations fongiques le brassage de l'information génétique semble principalement résulter de la reproduction sexuée, toutefois une parasexualité ainsi que des transferts horizontaux peuvent aussi être envisagés.

1.2- Absence de structure génétique au sein de population en panmixie

Dans une population théorique en panmixie, les croisements entre individus ont lieu de façon aléatoire et chaque individu de la population a autant de chance de se reproduire avec tout autre individu de la même population. La génération suivante est donc issue d'un mélange aléatoire de l'information génétique présente dans la population initiale. Les individus au sein d'une population panmictique sont donc issus du brassage génétique du même pool de gènes et aucune structure génétique ne peut être identifiée (Johansson et al., 2018). Les populations panmictiques sont des populations modèles, dans les populations naturelles il existe toujours des facteurs qui influent sur l'appariement des isolats qui ne sont alors jamais complètement aléatoires.

1.3- Différenciation de populations : le continuum de la spéciation

Quand on s'écarte d'une situation panmictique, la reproduction n'a plus lieu complètement au hasard au sein de la population. Des barrières au flux de gènes limitent la reproduction entre certains groupes d'individus. Ces barrières au flux de gènes peuvent être séparées en deux groupes, les barrières pré-appariement ou les barrières post-appariement qui bloquent respectivement le flux de gènes avant ou après la fusion des cellules. Parmi les barrières pré-appariement, Giraud et al. (2008) citent 5 types de

barrières. (1) Pour des espèces dépendantes de vecteurs biotiques, la spécialisation des vecteurs en eux même peut empêcher le contact entre deux populations. On peut citer par exemple *Microbotryum violaceum* dont les insectes vecteurs varient en fonction de l'espèce de la plante hôte. Ainsi des isolats présents sur différentes espèces hôtes ont moins de chances de se rencontrer (van Putten et al., 2007). (2) La spécialisation à des facteurs environnementaux peut également jouer ce rôle si la reproduction a lieu au sein de ces environnements (Giraud, 2006). (3) L'allochronie : reproduction à des instants différents. Ce type de barrières existe par exemple entre les espèces *Saccharomyces cerevisiae* et *S. paradoxus* (Murphy et al., 2006). (4) Un fort taux d'auto-fertilisation est aussi une barrière observée dans la nature (Giraud et al., 2008b). (5) L'appariement assortatif existe s'il y a un choix du partenaire sexuel empêchant ainsi certains croisements d'avoir lieu (Treindl and Leuchtmann, 2019). Les barrières post-appariement quant à elles regroupent des barrières associées à une non-viabilité ou une stérilité de la descendance (Giraud et al., 2008a).

Les différences génétiques existant entre des groupes d'individus séparés par une barrière au flux de gènes tendent alors à s'accroitre. Cette différenciation traduit une évolution plus ou moins indépendante des populations (selon la perméabilité de la barrière au flux de gènes) sous l'action de la dérive génétique ou de pressions de sélection. L'établissement des barrières aux flux de gènes, et leur perméabilité, est une question centrale en biologie évolutive pour comprendre l'apparition de la diversité du monde vivant et particulièrement l'apparition des espèces. Dès 1859, Darwin dans ses travaux émet l'hypothèse que les mécanismes entrainant des différenciations au sein des populations sont les mêmes mécanismes qui à terme entrainent la séparation des espèces :
« Les petites différences qui distinguent les variétés d'une même espèce tendent régulièrement à s'accroître jusqu'à ce qu'elles deviennent aux grandes différences qui existent entre les espèces d'un même genre, ou même de genres distincts. » (Darwin, 1859)

La spéciation est alors décrite comme un processus continu. La divergence génétique ou phénotypique entre deux populations peut varier de façon graduelle (Drès and Mallet, 2002; Jiggins and Mallet, 2000; Mallet et al., 2007; Nosil, 2007; Rymer et al., 2010; Thorpe et al., 2010), de même que l'intensité de l'isolement reproducteur (Coyne and Orr, 1989; Funk et al., 2006; Merrill et al., 2011; Nosil, 2007). Ce continuum de la spéciation est illustré par la Figure 1. Entre les deux extrêmes du processus de spéciation, à savoir d'une part une population unique non structurée et d'autre part deux espèces distinctes, De Queiroz (2007) définit une « zone grise » de la spéciation dans la laquelle certains aspects sont différenciés entre les populations tandis que d'autres ne le sont pas rendant difficile de conclure quant au nombre d'espèces (Figure 1).

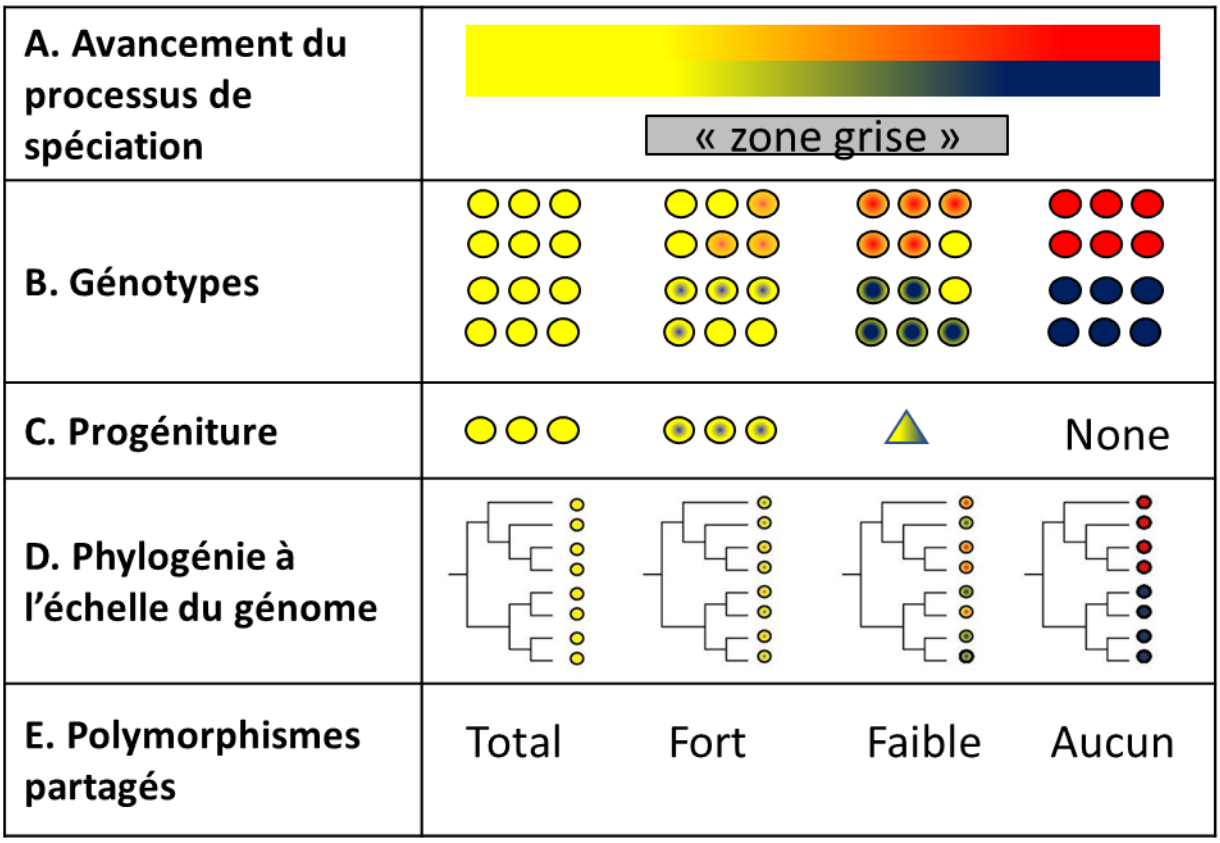


Figure 1: Représentation simplifiée du processus de spéciation

(Adapté de Matute and Sepúlveda, 2019)

A : Avancement dans le processus de spéciation. Cet avancement est continu rendant la spéciation parfois difficile à détecter dans les cas de divergence récente (zone grise de la spéciation) ; B : Au cours du processus les fréquences alléliques vont se différencier entre les deux populations pouvant s'accompagner d'une différenciation phénotypique (mais ce n'est pas toujours le cas) ; C : L'accroissement de l'isolement reproducteur est attendu avec l'accroissement de la divergence entre populations; D : La divergence des populations s'accompagne d'une accumulation de polymorphismes spécifiques à chaque population augmentant la probabilité que chacune des populations devienne monophylétique (i.e. qu'il y ait un « tri » des lignées); E : le nombre de polymorphismes partagés entre les populations décroit au cours du processus.

1.4- Les barrières aux flux de gènes initiant la spéciation

Les modes de spéciation peuvent être classés selon les mécanismes à l'origine de la divergence des populations (Nosil, 2012; Schluter, 2001). Ainsi trois modes de spéciation majeurs sont identifiés.

- Spéciation écologique

Une spéciation écologique a lieu dans les cas où un isolement reproducteur se met en place entre deux populations en réponse à des pressions de sélection écologiques et divergentes (Rundle and Nosil, 2005). Une pression de sélection est considérée comme écologique si elle est une conséquence de l'interaction entre des individus et leur environnement (Rundle and Nosil, 2005). L'environnement est ici considéré au sens large et inclut les éléments biotiques et abiotiques du milieu ainsi que les interactions intra et interspécifiques (compétition, prédation, ...). Afin que ces pressions de sélection divergentes impactent la structure des populations et conduisent à la spéciation, trois composantes sont nécessaires : (1) des pressions de sélection divergentes exercées sur une population ; (2) une forme d'isolement reproducteur ; (3) un mécanisme génétique liant l'isolement reproducteur à la sélection divergente (Rundle and Nosil, 2005).

- « Mutations-order speciation »

Un isolement reproducteur peut apparaître par la fixation au sein d'une même population initiale de différentes mutations avantageuses, incompatibles entre elles, face à une pression de sélection uniforme (Schluter, 2009). L'écologie peut être à la base des pressions de sélection exercées sur les populations, toutefois dans ce cas de figure ces facteurs écologiques ne favorisent pas la divergence (la pression de sélection étant uniforme, les mêmes phénotypes sont favorisés dans les deux populations). Dans ce modèle, aucune corrélation n'est attendue entre isolement reproducteur et divergence écologique (Nosil, 2012).

- Spéciation sans sélection

Ce mode de spéciation ne dépend pas de pressions de sélection (toutefois des pressions de sélection peuvent s'appliquer par la suite). Il englobe les cas d'apparition de barrières physiques entre populations suivis par une dérive génétique de celles-ci, ainsi que les cas d'effet fondateur ou goulets d'étranglement des populations. La dérive génétique est un mécanisme évolutif qui consiste en la modification des fréquences alléliques entre générations causée par le hasard (échantillonnage aléatoire de lignées au sein du pool génétique constituant la génération précédente). Ces cas de figure entrainent évidemment une structuration des populations par l'apparition d'importantes barrières au flux de gènes. Toutefois, l'apparition d'un isolement reproducteur entre deux populations par le simple

effet de la dérive génétique et/ou d'un effet fondateur est débattue dans la communauté scientifique (Coyne and Orr, 2004; Rundell and Price, 2009; Templeton, 2008). Coyne and Orr (2004) concluent qu'une spéciation complète par ces seules barrières serait en théorie difficile et rare dans la nature.

Certains cas d'hybridation peuvent entrainer des modifications du nombre de chromosomes (polyploïdie ou aneuploïdie) rendant impossible une reproduction sexuée entre ces individus et les individus de la population initiale et créant un isolement reproducteur immédiat entre ces populations (Giraud et al., 2008a).

1.5- Structuration des populations et spéciation chez les champignons phytopathogènes

Certaines particularités biologiques des champignons phytopathogènes ont un impact majeur sur la structure des populations au sein de ces espèces et sur la formation de nouvelles espèces.

1.5.1- Importance de l'hôte sur la structure des populations fongiques phytopathogènes

Comme mentionné ci-dessus, une condition nécessaire pour qu'un processus de spéciation écologique puisse avoir lieu, est l'établissement d'une liaison génétique entre les gènes sous sélection divergente et les gènes responsables de l'isolement reproducteur. L'apparition d'une mutation qui permettrait un avantage sélectif tout en ayant un effet pléiotrope sur le choix des partenaires sexuels a été qualifié de « trait magique » par Gavrilets (2004). Giraud et al. (2010) défendent que le cycle de vie des champignons phytopathogènes permet des phénomènes de spéciations écologiques rapides, particulièrement via des adaptations à de nouveaux hôtes, et estiment que les émergences de nouvelles maladies fongiques phytopathogènes sont souvent dues à un évènement de spéciation engendré par un saut d'hôte. En effet, chez les champignons phytopathogènes l'infection d'un nouvel hôte requiert en général la présence de mutations génétiques permettant d'éviter sa reconnaissance par les systèmes de défense de la plante. Seuls les individus adaptés à cet environnement (le nouvel hôte) auront la capacité de se développer sur celui-ci et d'atteindre le stade du cycle de vie permettant la reproduction sexuée. Or, la plupart des espèces fongiques ne sont pas mobiles durant cette phase de croissance. Ainsi, la reproduction sexuée ne peut avoir lieu qu'entre individus adaptés à la même espèce hôte et de ce fait, toute mutation permettant l'infection d'un nouvel hôte a un effet pléiotrope sur l'isolement reproducteur et peut être considéré comme un trait magique (Giraud et al., 2010).

1.5.2- Importance du mode de reproduction sur la structure des populations fongiques phytopathogènes

La plupart des espèces fongiques peuvent se reproduire à la fois de façon sexuée et asexuée. L'alternance entre les deux cycles est très variable d'une espèce fongique à l'autre. Certaines espèces

nécessitent une phase de reproduction sexuée (ce qui peut être le cas par exemple quand la reproduction sexuée permet la survie dans des conditions environnementales défavorables), chez d'autres espèces la reproduction sexuée semble n'avoir lieu qu'en de très rares occasions et enfin, certaines espèces, telles que *Pseudocercospora fijensis*, *Fusarium graminearum* ou *Phyllostica citricarpa*, se reproduisent sexuellement et asexuellement tout au long de l'année. Le mode de reproduction influe fortement sur la structure des populations. Une population strictement asexuée ou autogame est par définition isolée des autres populations et évolue de façon indépendante. Toutefois, le statut d'espèce des populations strictement asexuées ou autogames est débattu. Plus aucun brassage génétique n'ayant lieu au sein de ces populations, tous les individus évoluent de façon indépendante les uns des autres. Coyne and Orr (2004) considèrent alors qu'une population clonale ne constitue pas une espèce dans sa globalité mais un ensemble de « micro-espèces » composées d'individus propageant chacun leur propre lignée génétiquement isolée des autres. Toutefois, il semble que les organismes asexués forment souvent en conditions naturelles des groupes discrets d'individus, et non une distribution continue de phénotypes/génotypes comme on pourrait s'y attendre, pouvant être considérés par certains comme des espèces différenciées (Fontaneto and Barraclough, 2015; Fontaneto et al., 2007). Les hypothèses mises en avant pour expliquer ces profils sont l'existence de niches écologiques discrètes, l'extinction de certains des génotypes/phénotypes intermédiaires ou de multiples apparitions de populations clonales à partir d'une population sexuée (Coyne and Orr, 2004; Giraud et al., 2008b).

2- Concept d'espèce et méthode de reconnaissances des espèces chez les champignons

2.1- Concept d'espèce

Définir une espèce n'est pas trivial. Ce sujet a été beaucoup étudié, débattu et adapté avec l'avancée de nos connaissances à propos de l'évolution des organismes vivants (De Queiroz, 2007; Mallet et al., 2007). Dans une revue, De Queiroz (2007) ne décrit pas moins de 14 différents concepts d'espèces proposés par la communauté scientifique. Il explique ce grand nombre de définitions et l'absence de consensus par la confusion existant entre « concept d'espèce » et « critères de définition des espèces ». De l'ensemble de ces propositions, il détermine une idée commune qui permettrait de définir les espèces comme des « segments de lignées évolutives qui évoluent indépendamment les unes des autres » (De Queiroz, 2007; Giraud et al., 2008a). Il propose alors de faire le distinguo entre le concept d'espèce qui serait la définition consensus définie ci-dessus et les critères de reconnaissance d'espèces permettant de trancher si deux individus appartiennent ou non à la même espèce.

Tableau 1 : Critères de reconnaissance d'espèces décrits dans la littérature (adaptée de De Queiroz, 2007)

Critère de reconnaissance d'espèce	Propriétés	Références
Isolement	Isolement reproducteur intrinsèque (absence de croisements entre organismes hétérospécifiques sur la base de propriétés intrinsèques, par opposition à des barrières extrinsèques, qui dépendent de l'écologie ou la distance géographique)	*(Dobzhansky, 1950; Mayr, 1999)*
Reconnaissance	Système commun de reconnaissance ou de fécondation du partenaire (mécanismes par lesquels des organismes conspécifiques, ou leurs gamètes, se reconnaissent mutuellement pour l'accouplement et la fécondation)	*(Lambert and Spencer, 1996; M. L. Masters, 1987; Paterson, 1985)*
Ecologie	Même niche écologique (tous les composants de l'environnement avec lesquels des organismes conspécifiques interagissent)	*(Andersson, 1990; Valen, 1976)*
Monophylie	Monophylie (composée d'un ancêtre et de tous ses descendants ; généralement déduit de la possession d'états de caractères dérivés partagés)	*(Donoghue, 1985; Mishler, 1985; Rosen, 1979)*
Généalogie	Coalescence exclusive d'allèles (tous les allèles d'un gène donné descendent d'un allèle ancestral commun non partagé avec ceux d'autres espèces)	*(Avise and Ball, 1990; Baum and Shaw, 1995)*
Diagnostic	Diagnosticabilité (différence qualitative, fixée)	*(Cracraft, 1983; Nelson and Platnick, 1982; Nixon and Wheeler, 1990)*
Phénétique	Forme un cluster phénétique (différence quantitative)	*(Michener, 1970; Sneath and Sokal, 1973, 1973)*
Cluster Génotypique	Forme un cluster génotypique (déficits en génotypes intermédiaires ; par exemple, hétérozygotes)	*(Mallet, 1995)*

2.2- Critères de reconnaissance d'espèces

La difficulté de définir des critères de reconnaissance d'espèces vient certainement de trois raisons principales (Giraud et al., 2008a) :

- La spéciation est un processus continu qui progresse de façon très différente d'une espèce à l'autre ;
- La spéciation peut être due à un grand nombre de facteurs différents ce qui influe sur l'ordre d'apparition des caractères pouvant être utilisés pour la distinction des espèces ;
- Les caractéristiques biologiques de certains organismes rendent certains critères difficiles à évaluer.

Ainsi la plupart des multiples « concepts d'espèce » proposés dans la littérature représentent plutôt différents critères de reconnaissance des espèces. En effet, comme présenté précédemment plusieurs modifications sont attendues lors du processus de spéciation (modification des fréquences alléliques, mise en place d'un isolement reproducteur, tri des lignées, différenciation morphologique …). Parmi ces critères de reconnaissances d'espèces, les plus connus sont décrits ci-dessous. Une liste plus exhaustive est proposée dans le Tableau 1.

Morphological species recognition (MSR) : Ce critère de reconnaissance d'espèces se base sur des différences morphologiques entre espèces. Il est resté jusqu'à récemment le critère le plus utilisé pour la distinction d'espèces fongiques (Giraud et al., 2008a). Plus de 70,000 espèces fongiques ont été décrites sur la base de critères morphologiques ou autres caractères phénotypiques tels que la vitesse de croissance à différentes températures, la production de métabolites secondaires ou la présence de pigments (Besl and Bresinsky, 1997; Frisvad and Filtenborg, 1990; Pitt, 1979; Taylor et al., 2000). Toutefois, de nombreuses espèces ne présentent pas de différences morphologiques entre elles et ce critère de reconnaissance ne permet pas leur distinction (Giraud et al., 2008a; Taylor et al., 2000).

Biological species recognition (BSR) : Ce critère se base sur l'apparition d'un isolement reproducteur entre espèces. Deux espèces sont ainsi distinguées si elles sont dans l'incapacité de produire une descendance viable et fertile. Ce critère a parfois permis de distinguer des espèces morphologiquement identiques. On peut citer par exemple l'espèce de champignon basidiomycète définie morphologiquement, *Armillaria mellea sensu lato*, au sein de laquelle une douzaine d'espèces ont pu être décrites par des tests de reproduction (Anderson and Ullrich, 1979; Qin et al., 2007). Ce critère ne peut s'appliquer pour des organismes pour lesquels on ne peut pas induire de reproduction sexuée au laboratoire.

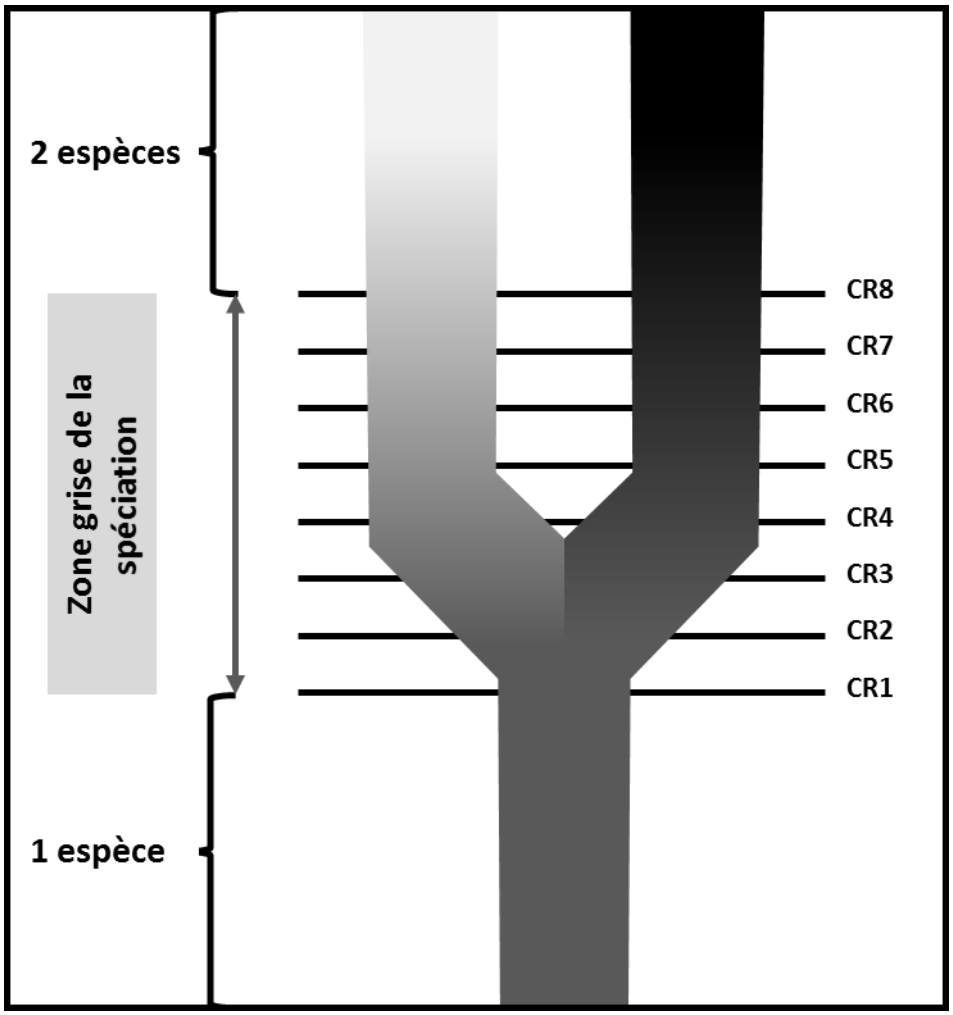


Figure 2: Schématisation des apparitions successives des différents critères permettant la distinction d'espèces au cours du processus de spéciation.

(Adapté de De Queiroz, 2007)

Le dégradé de gris représente la différenciation au cours du temps de deux lignées (deux espèces) à partir d'une même lignée (espèce) initiale. Les lignes horizontales nommées CR1 à CR8 représentent le moment ou les deux lignées se distinguent selon un critère de reconnaissance d'espèce donné (morphologiquement différentes, sexuellement incompatibles, écologiquement distinctes…). La zone grise de la spéciation correspond à la période ou deux lignées peuvent être considérées comme deux espèces distinctes selon certains critères mais pas selon l'ensemble des critères (critères théoriques nommés arbitrairement CR1 à CR8 sur la figure).

Ecological species recognition (ESR) : Ce critère se base sur l'adaptation des espèces à des niches écologiques différentes. Ce critère a été proposé pour expliquer les profils complexes observés au sein du genre *Quercus* (Burger, 1975; Valen, 1976).

Phylogenetic species recognition (PSR) : Ce critère se base sur la divergence nucléotidique entre espèces. Une extension de ce critère est le « **Genealogical Concordance Phylogenetic Species Recognition (GCPSR)** » qui implique la concordance phylogénétique de multiples gènes non liés entre eux. Ce dernier critère a été l'un des plus utilisé concernant les espèces fongiques (Dettman et al., 2003; Fournier et al., 2005; Giraud et al., 2008a; Johnson et al., 2005; Koufopanou et al., 2001; Le Gac et al., 2007). L'étude de la divergence nucléotidique permet en général des différenciations plus résolutives des espèces ayant récemment divergé.

Pourtant les différents marqueurs de la spéciation ne divergent pas tous en même temps et au même rythme (De Queiroz, 2007; Roux et al., 2016). Par exemple, des différenciations morphologiques claires peuvent être visibles entre deux populations sans toutefois que les lignées soient réciproquement monophylétiques (Dettman et al., 2007). Dans d'autres cas de figures, les différences morphologiques sont inexistantes entre deux populations fortement différenciées génétiquement (Amato et al., 2007). Ce deuxième cas de figure est toutefois le plus fréquent concernant les espèces fongiques (Taylor et al., 2006). Ainsi, l'étude de chacun de ces critères de façon indépendante n'aboutit pas toujours aux mêmes conclusions sur le statut d'espèce des populations étudiées (De Queiroz, 2007). L'étude des divergences nucléotidiques a toutefois le meilleur potentiel pour la détection de divergences récentes. La divergence successive dans le temps de chacun des critères de reconnaissance des espèces est illustrée dans la Figure 2.

2.3- Méthodes de détection opérationnelles d'organismes pathogènes

Les maladies fongiques des plantes peuvent représenter une forte menace pour la sécurité alimentaire ainsi que pour l'économie agricole (Dean et al., 2012; Doehlemann et al., 2017). Un grand nombre de pays ont ainsi instauré des contrôles sanitaires à leurs frontières afin de prévenir l'introduction d'agents phytopathogènes (Mancini et al., 2016). Les listes d'organismes de quarantaine sont basées sur des dénominations taxonomiques. La plupart des agents pathogènes listés sont identifiés au niveau de l'espèce, toutefois des genres, formae speciales ou encore « isolats pathogènes non européens » peuvent également être listés. Il est alors essentiel de connaitre le niveau taxonomique ciblé, et la structure génétique de l'agent pathogène afin de pouvoir mettre en place des techniques de détection adaptées.

Tableau 2 : Principales techniques utilisées pour le diagnostic de champignons et leurs principales caractéristiques

TECHNIQUES	CARACTERISTIQUES	AVANTAGES	DESAVANTAGES	EXEMPLES
Examen visuel	Examen visuel de symptômes et identification de l'agent pathogène par ses caractéristiques morphologiques (symptômes, mycélium, spores)	- Ne nécessite pas de connaissances génétiques	- Détection impossible pendant les phases asymptomatiques - Imprécision de l'identification (ne permet pas d'identification en dessous du niveau de l'espèce, voire du genre) - Nécessite une bonne expertise taxonomique	(Murakishi, 1951; Randall-Schadel et al., 2001)
Incubation	Incubation sur des milieux sélectifs	- Simple d'utilisation	- Imprécision de l'identification - Nécessite une bonne expertise taxonomique	(Garibaldi et al., 2004)
Tests sérologiques	Développement et détection d'anticorps ciblant l'agent pathogène	- Ne requiert pas de culture pure	- Peu développés pour les champignons - Détecte les agents pathogènes non-viables	(Afouda et al., 2009)
Amplification d'ADN		- Précision de l'identification	- Détecte les agents pathogènes non-viables - Peut être inhibé par des composés présent dans les matrices végétales	(Ward et al., 2004)
	PCR	Amplification d'ADN basée sur la spécificité des amorces		- Test non-quantitatif	(Kuzdraliński et al., 2017)
	qPCR	Amplification d'ADN basée sur la spécificité des amorces et suivie par l'émission d'une fluorescence	- Test quantitatif - Grande sensibilité	- Coûteux	(Schena et al., 2004)
	Nested PCR	Deux réactions de PCR emboitées	- Grande sensibilité	- Coûteux en prix et en temps - Sensible aux contaminations	(Chiocchetti et al., 2001)
	LAMP	Amplification isothermale	- Rapidité des résultats - Peu couteux en réactifs et matériel - Utilisable sur le terrain - Grande sensibilité	- Sensible aux contaminations - difficulté de conception (spécificité délicate à obtenir)	(Mumford et al., 2006)
	PCR-luminex	Hybridation entre un produit de PCR biotinylé et une sonde nucléotidique couplée à des microbilles permettant l'émission d'une fluorescence	- Important multiplexage possible	- Coût de l'équipement dédié	(Ishii et al., 2008)
NGS	Séquençage de génomes complet ou partiel (incluant l'inventaire des communautés par metabarcoding)	-Précision de l'identification	- Peu utilisé pour la détection de champignon - Délai d'analyse, peu adapté à la détection en routine	(Nicolaisen et al., 2014)

Quarante-neuf taxons fongiques sont ainsi réglementés par l'Union Européenne, listés dans des annexes de directives comme agents pathogènes de quarantaine et sont soumis à des contrôles aux frontières (Council Directive 2000/29/EC, Annex Ia and IIa). Cette directive sera toutefois abrogée et remplacée par le règlement (UE) 2016/2031 relatif aux mesures de protection contre les ennemis des végétaux qui sera applicable à partir du 14 décembre 2019.

Les contrôles sont généralement basés sur un échantillonnage des végétaux ciblé visuellement (en présence de symptômes) ou déterminé statistiquement (en absence de symptômes, par exemple sur semences), puis sur l'utilisation de tests de détection des agents pathogènes recherchés. La fiabilité de ces tests est essentielle car tout diagnostic erroné peut avoir de lourdes conséquences. Un organisme pathogène qui ne serait pas détecté pourrait être introduit dans de nouvelles régions tandis que des résultats faussement positifs pourrait entrainer la destruction de matériel végétal sain. Ces tests de détection doivent alors remplir certaines conditions (Broeders et al., 2014; Lievens and Thomma, 2005):

- Inclusivité : capacité de produire un résultat positif pour tous les individus ciblés, malgré l'éventuelle variabilité génétique présente au sein du taxon recherché ;
- Spécificité : Capacité de produire un résultat négatif pour tous les individus non-cibles ;
- Sensibilité : Quantité minimale de taxon-cible détectable grâce au test ;
- Robustesse : Fiabilité des résultats (sensibilité, spécificité) en cas de légers écarts au protocole d'utilisation du test ;
- Transférabilité : Fiabilité des résultats du test lors de son utilisation dans des conditions expérimentales différentes de celles initialement décrites pour son développement (différents laboratoires, matériel, manipulateurs ou réactifs).

Différentes techniques peuvent être utilisées pour la détection de champignons phytopathogènes (Crous et al., 2015; Mancini et al., 2016). Un certain nombre de ces techniques ainsi que leurs principales caractéristiques sont présentées dans le Tableau 2.

Figure 3 : Illustration des symptômes causés par *P. oryzae* sur riz

A : Symptômes caractéristiques de forme elliptique causés par *P. oryzae* sur une feuille de riz collectée au champ (photo de Jean-Loup Nottéghem - CIRAD) ; B : Nécrose du cou d'une panicule de riz entrainant la nécrose totale de la panicule (photo de Jean-Loup Nottéghem - CIRAD) ; C : Parcelles expérimentales en Camargue avec, à gauche une variété de riz sensible fortement touchée par la pyriculariose, et à droite une variété résistante (photo de Jean-Benoit Morel - INRA)

3- Diversité de *Pyricularia oryzae* et de ses hôtes

3.1- Impact et symptômes de la pyriculariose

La pyriculariose est une maladie qui touche un grand nombre de poacées dont le riz et le blé. Les premiers écrits faisant référence à la pyriculariose datent de 1637 en Chine où elle a été observée dans des champs de riz. Aujourd'hui on estime à plus de 4% les pertes de rendement dues à cette maladie, sur les cultures de riz, à l'échelle mondiale (Savary et al., 2019). Les symptômes décrits sont des nécroses sur les parties aériennes des plantes infectées : tiges, feuilles, panicule et grains (Ou, 1985). Les symptômes observés sur les feuilles sont parmi les plus caractéristiques. Sur une variété sensible, les lésions ont une forme elliptique (« œil de chat »). Leur centre se teinte d'une couleur grisâtre voire blanchâtre lors de la sporulation du champignon tandis que les bordures de la lésion sont généralement de couleur marron (*Figure 3*). Ce sont toutefois les symptômes paniculaires qui impactent le plus les rendements et la qualité des grains. Dans ce type d'infection, le champignon entraine une nécrose au niveau du racème ou du cou (nœud situé directement sous le racème) empêchant le remplissage des grains (*Figure 3*). La gravité des symptômes varie en fonction du stade de l'infection, des conditions environnementales, de la quantité d'inoculum ou du niveau de résistance de la plante hôte.

3.2- Cycle infectieux de l'agent pathogène

L'agent causal de la pyriculariose est le champignon ascomycète hémibiotrophe *Pyricularia oryzae* (Syn. *Magnaporthe oryzae*). Son cycle infectieux commence par l'arrivée d'une spore sur la plante hôte. Au contact de la plante et avec des conditions d'humidité suffisante, la spore peut adhérer à la surface de la plante hôte. Trente minutes d'hydratation sont nécessaires pour l'adhésion des spores (Ikeda et al., 2019). La germination des spores est suivie par la formation d'un appressorium, organe permettant l'entrée des hyphes dans les tissus de la plante en exerçant une forte pression et en perforant la paroi. Les premiers jours d'infection correspondent à une phase biotrophe asymptomatique (le champignon se développe sans tuer les tissus foliaires). Tout au long de ce processus infectieux, le champignon sécrète de petites protéines (effecteurs) qui inactivent les réactions de défense de la plante et qui permettent que l'infection puisse suivre son cours. Au bout de 3-4 jours, l'agent pathogène passe à une phase nécrotrophe provoquant la nécrose des tissus de la plante hôte et menant aux symptômes décrits plus hauts.

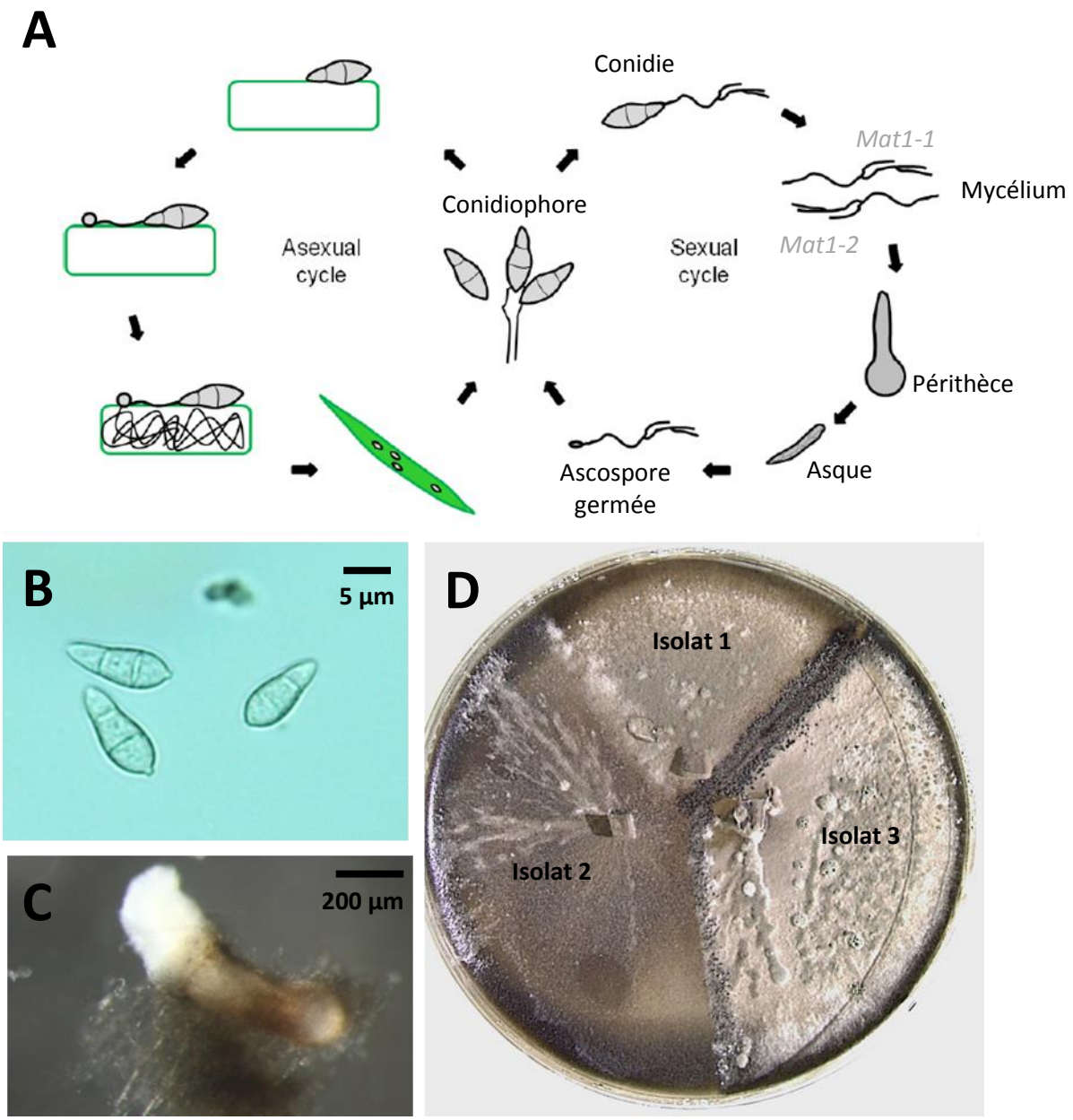


Figure 4 : Cycles de reproduction sexuée et asexuée de *P. oryzae*

A : représentation schématique des cycles de reproduction sexué et asexué de *P. oryzae (Adapté de la thèse de doctorat de Dounia Saleh)* ; B : conidies ; C : périthèce ; D : Formation de périthèces à la zone de contact du mycélium de trois isolats (pas de périthèces formés entre l'isolat 1 et l'isolat 2 ; périthèces formés uniquement par l'isolat 3 dans l'interaction entre l'isolat 2 et l'isolat 3 ; périthèces formés par les deux isolats dans l'interaction de l'isolat 1 et l'isolat 3).

3.3- Cycle de reproduction de *Pyricularia oryzae*

Le cycle de reproduction de *Pyricularia oryzae* n'a été étudié quasi exclusivement que sur les populations du champignon infectant le riz. Sur cette céréale, il a été montré que cette espèce fongique, comme un grand nombre d'espèces fongiques, est capable à la fois de reproduction sexuée et asexuée (*Figure 4*). Lors du cycle de reproduction asexué, des conidiospores (ou conidies) sont formés à l'extrémité de conidiophores. Une à 20 conidies peuvent être formées sur chaque conidiophore. Les conidies sont séparées par 2 septa et ont une forme de poire caractéristique à l'origine du nom *Pyricularia*. Le cycle de reproduction sexué de ce champignon hétérothallique ne peut avoir lieu qu'entre deux isolats présentant des types sexuels opposés (Mat1-1 et Mat1-2). Quand la reproduction sexuée a lieu entre deux individus, des fructifications, les périthèces, sont formés à la jonction entre les deux mycéliums par l'un, l'autre ou les deux isolats. Au sein des périthèces a lieu la fusion des gamètes, puis une méiose et une mitose permettent la formation d'un asque contenant 8 ascospores. Une autre condition est alors indispensable pour que la reproduction sexuée puisse avoir lieu : au moins un des deux isolats doit avoir la capacité à former des périthèces, c'est à dire être femelle-fertile. Cette fertilité femelle est souvent perdue chez les isolats dans les populations naturelles pathogènes du riz (Saleh et al., 2012).

Bien que le cycle de reproduction sexuée puisse être induit en condition de laboratoire, il est important de noter que jamais des structures de reproduction sexuée n'ont pu être observées en conditions naturelles. Toutefois, l'accumulation d'un certain nombre de preuves indirectes indique que certaines populations au sein de l'espèce ont recours à ce mode de reproduction (Kumar et al., 1999; Saleh et al., 2012; Zeigler, 1998). Une population en particulier, échantillonnée sur du riz dans le centre d'origine de l'agent pathogène (province chinoise du Yunnan), a été caractérisée par Saleh et al. en 2012. Au sein de cette population les deux types sexuels Mat1-1 et Mat1-2 ont été observés à des fréquences équivalentes. Lors de croisements en condition de laboratoire la plupart des isolats se sont révélés femelle-fertiles et les croisements compatibles ont abouti à la formation d'une descendance viable. Le faible déséquilibre de liaison ainsi que la forte richesse génotypique au sein de cette population impliquait également une reproduction sexuée. Afin de valider cette hypothèse, un nouvel échantillonnage a été fait l'année suivante. L'analyse des génotypes obtenus a mis en évidence un fort brassage allélique entre les deux années d'échantillonnage confirmant la présence d'une reproduction sexuée dans cette population. Dans d'autres régions du monde toutefois, les populations échantillonnées sur du riz ne semblent se reproduire que par reproduction asexuée (Choi et al., 2013; Correll et al., 2000; Gladieux et al., 2018a; Levy et al., 1993; Saleh et al., 2014; Thuan et al., 2006).

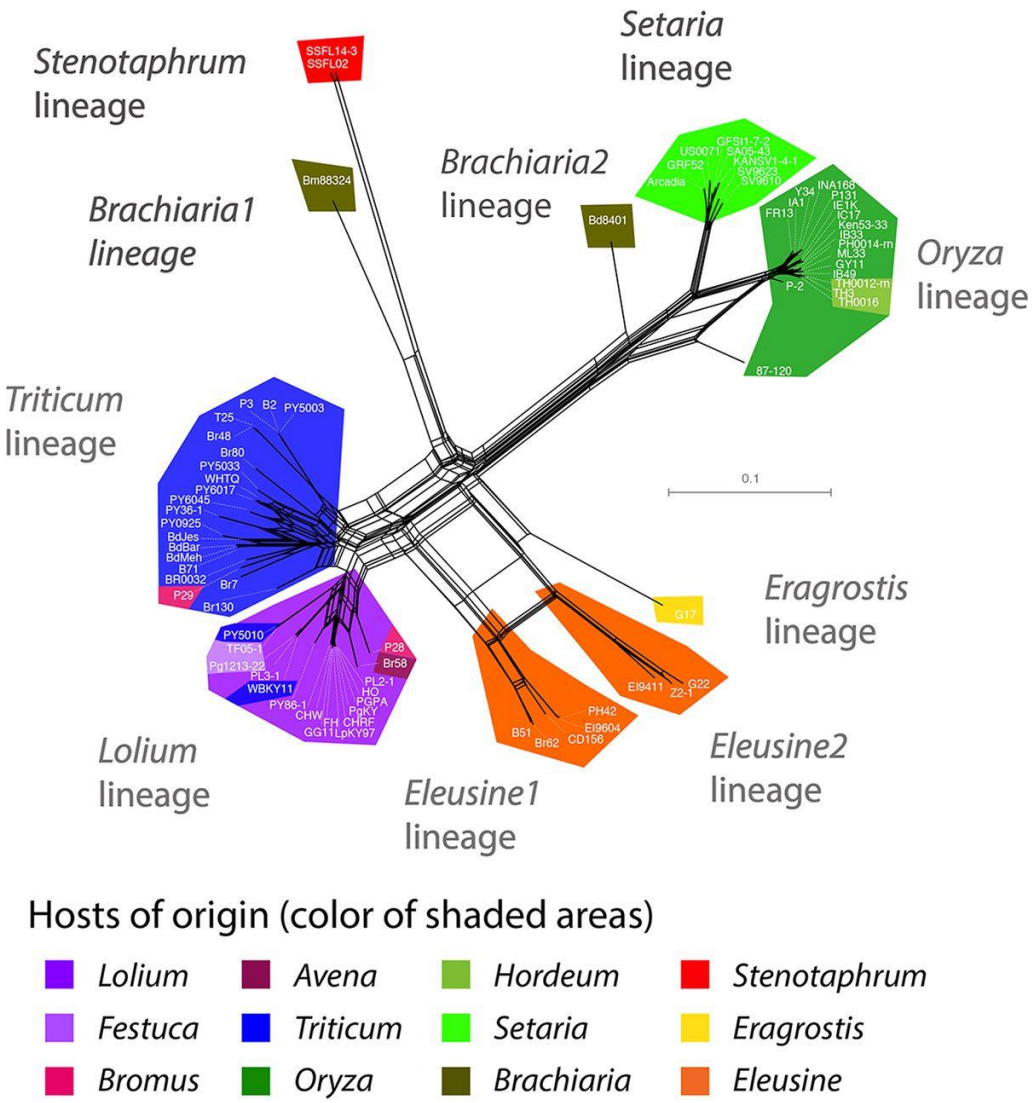


Figure 5 : Lignées hôte-spécifiques au sein de *P. oryzae*

(Extrait de Gladieux et al. 2018b)

Réseau phylogénétique, construit grâce au logiciel Splitstree, basé sur 25,078 SNP identifiés dans les séquences de 2,682 gènes orthologues en simple copie extraites des génomes de 76 isolats de P. oryzae isolés sur 12 genres d'hôtes différents.

3.4- Spectre d'hôte de *Pyricularia oryzae*

Les symptômes de la pyriculariose ont été décrits sur un grand nombre de poacées. Dans l'étude taxonomique menée par Klaubauf et al. (2014), 12 espèces hôtes de *P. oryzae* sont référencées. Pourtant, l'évaluation des spectres d'hôtes d'isolats de *P. oryzae* étudiés individuellement a montré que chaque isolat semble n'être pathogène que sur un seul, ou un nombre restreint, d'espèces hôtes (Couch et al., 2005; Kato et al., 2000; Tosa et al., 2016) et est généralement plus agressif sur son hôte d'origine (espèce sur laquelle il a été isolé). Ainsi, dès 1977 on utilise le terme de pathotype afin de pouvoir différencier les populations du champignon capables d'infecter l'éleusine, de populations infectant le riz (Valent et al., 2019). Des sauts d'hôte récents de l'agent pathogène ont pu être observés entrainant l'émergence de nouvelles maladies. Des épidémies sur des espèces de ray-grass sont apparues dans les années 90 aux USA, tandis que les premières épidémies affectant le blé se sont déclarées en 1985 au Brésil (Cruz and Valent, 2017).

3.5- Structure des populations au sein l'espèce *Pyricularia oryzae*

3.5.1- 1er niveau de structure : lignées hôte-spécifiques

L'étude de marqueurs génétiques a permis de mettre en évidence des lignées hôte-spécifiques au sein de l'espèce concordant avec les pathotypes précédemment identifiés (Couch et al., 2005; Gladieux et al., 2018b). En se basant sur 8000 gènes orthologues répartis dans l'ensemble du génome, Gladieux et al. (2018b) ont montré que les isolats capables d'infecter le riz sont fortement apparentés génétiquement et forment une lignée monophylétique nommée lignée Oryza. Dix lignées hôtes spécifiques sont ainsi décrites par Gladieux et al., (2018b) grâce à l'étude de génomes complets de 76 isolats de *P. oryzae* (*Figure 5*). La présence de ces lignées suppose une spéciation naissante suite à des sauts d'hôte ou à des expansions du spectre d'hôtes du champignon. Toutefois la divergence génétique entre ces lignées reste très faible (<1% de divergence) et un flux de gène et/ou un tri incomplet des lignées a été mis en évidence entre les lignées hôte-spécifiques (Gladieux et al., 2018b). Selon le critère d'espèce décrit par Dettman et al., (2003), le « Genealogical Concordance Phylogenetic Species Recognition », impliquant que le clade doit être fortement supporté par la généalogie d'au moins un locus et n'être contredit par aucune autre généalogie, les lignées hôte-spécifiques ne peuvent être considérées comme des espèces différentes (Gladieux et al., 2018b). Ces lignées hôte-spécifiques sont toutefois responsables de maladie distinctes ne touchant pas les mêmes espèces-hôtes et n'ayant pas les mêmes aires de répartition. La lignée Oryza par exemple infecte le riz est présente dans le monde entier dans toutes les zones de production de cette céréale (Ou, 1980), tandis que la lignée Triticum est responsable de la pyriculariose du blé, maladie émergente circonscrite à l'Amérique du sud et à

deux pays d'Asie (Islam et al., 2019). La mise au point d'outils de diagnostic spécifiques à chacune de ces lignées est donc cruciale d'un point de vue sanitaire malgré leur forte proximité génétique, et particulièrement en ce qui concerne la lignée Triticum dont la propagation menace les cultures de blé mondiales.

3.5.2- 2ème niveau de structure au sein de la lignée Oryza

La lignée Oryza infectant le riz est de très loin la lignée la plus étudiée au sein de *P. oryzae*. En effet, le riz est la source nutritionnelle majoritaire pour la moitié de la population mondiale et les dégâts causés par *P. oryzae* sur cette céréale ont été estimés à plus de 4% de perte du rendement mondial (Dean et al., 2012; Khush, 2005; Savary et al., 2019). De plus, la pyriculariose du riz s'est propagée mondialement et est présente dans l'ensemble des pays producteurs de cette céréale (Ou, 1980).

Afin de comprendre l'évolution de l'agent pathogène et d'adapter les techniques de lutte et les programmes d'amélioration variétale, la structure génétique des populations de *P. oryzae* infectant le riz a été très étudiée. De multiples études à des échelles locales ont mis en évidence l'existence de structures génétiques au sein de cette lignée identifiant un nombre très variable de « sous-lignées » (pouvant aller de 2 lignées à 56) selon le lieu de l'expérimentation, le moment de l'expérimentation ou les techniques utilisés pour obtenir des marqueurs génétiques (Don et al., 1999; Kumar et al., 1999). Parmi ces techniques on peut citer la « Random Amplification of Polymorphic DNA » (RAPD), la « Repetitive element palindromic PCR » (rep-PCR), l'analyse de marqueurs microsatellites, le « génotype by sequencing » (GBS), des « Amplified fragment length polymorphism » (AFLP) ou le séquençage de génomes complets. La multiplicité des techniques employées et le manque d'isolats de référence rendent difficile la comparaison des lignées identifiées entre localités.

Certaines particularités, liées principalement au mode de reproduction, sont toutefois comparables entre ces populations. La plupart des populations sont considérées comme clonales car un grand déséquilibre de liaison est mis en évidence entre les marqueurs génétiques étudiés (Choi et al., 2013; Kumar et al., 1999; Thuan et al., 2006). Certaines populations étudiées ne présentent même qu'un seul type sexuel empêchant toute chance de reproduction sexuée (Mat1-1 : Argentine (Consolo et al., 2005), Iran (Hemmati et al., 2005), Corée (Park et al., 2003, 2008); Mat1-2 : USA (Pagliaccia et al., 2018), Vietnam (Thuan et al., 2006)). Une potentielle population ayant recours à la reproduction sexuée a tout de même été identifiée en 1999 dans une région de l'Inde himalayenne (Kumar et al., 1999). L'existence d'une reproduction sexuée a été par la suite confirmée dans une population Chinoise du Yunnan, également dans la région Himalayenne (Saleh et al., 2012). Cette région est considérée comme étant le centre d'origine et de diversité de la lignée Oryza de *P. oryzae* (Saleh et al., 2014).

Des études de la structure des populations effectuées à l'échelle globale, avec un échantillonnage international des isolats, s'accordent toutefois toutes sur un faible nombre de lignées allant de 3 à 6 lignées identifiées à travers le monde (Gladieux et al., 2018a; Saleh et al., 2014; Tharreau et al., 2009; Zhong et al., 2018). J'ai pu contribuer à l'une de ces études, portée par Pierre Gladieux (2018a), celle-ci est présentée à la fin de cette introduction. Dans chacune de ces études, une des lignées identifiées présentait un ratio assez équilibré des types sexuels et un signal de recombinaison (quand cela a pu être mesuré) compatible avec un mode de reproduction sexué (Gladieux et al., 2018a; Saleh et al., 2014; Tharreau et al., 2009; Zhong et al., 2018). Les isolats appartenant à cette lignée recombinante ont principalement été échantillonnés en Asie du sud-est. Les autres lignées identifiées sont considérées comme clonales car ne comprennent qu'un seul type sexuel (Mat1-1 ou Mat1-2 selon la lignée) et ne présentent pas de signal de recombinaison. Les aires de répartition de ces lignées sont variables, certaines semblent pandémiques tandis que d'autres semblent restreintes au continent asiatique (Gladieux et al., 2018a; Saleh et al., 2014). Le continent européen se distingue par la présence d'une seule lignée (Gladieux et al., 2018a; Saleh et al., 2014). La divergence de ces lignées a été datée à environ un millier d'années par deux études indépendantes (Gladieux et al., 2018a; Zhong et al., 2018). Des flux de gènes vers la lignée recombinante ont pu être identifiés par la détection dans la lignée recombinante de marqueurs génétiques spécifiques des lignées clonales (Gladieux et al., 2018a). Pour finir, il existe également une structure génétique au sein de l'hôte, *Oryza sativa,* au sein de laquelle plusieurs groupes génétiques sont identifiés parmi lesquels les groupes Japonica tempéré, Japonica tropical, Indica, Aus et Aromatic (Garris et al., 2005). Gladieux et al., (2018a) mettent en évidence une corrélation entre certaines lignées clonales au sein de *P. oryzae* et certains types de riz. Ainsi, les isolats appartenant à l'une des lignées clonales pandémiques ont principalement été échantillonnés sur des riz Japonica tandis qu'une autre lignée est quant à elle associée à des riz Indica. La structure au sein de la lignée oryza ainsi que la connaissance des facteurs impactant cette structure doivent être affinées pour mieux comprendre l'évolution et la propagation de l'agent pathogène à l'échelle globale.

4- Objectifs de la thèse

Dans le cadre de cette thèse nous nous sommes intéressés de façon indépendante aux deux niveaux de structure observés au sein de *P. oryzae.*

La structuration en lignées hôte-spécifiques a permis la mise en évidence de la lignée Triticum regroupant les isolats responsables de la pyriculariose du blé, maladie émergente dévastatrice. Le regroupement des isolats responsables de cette maladie au sein d'une même lignée implique la présence de régions génomiques spécifiques de ces isolats qui pourraient être ciblées pour la mise au point d'outils de d'outils de diagnostic moléculaires essentiels pour la gestion des épidémies de pyriculariose du blé. Sur la base de cette structure génétique, l'un des objectifs de la thèse était d'identifier par une approche de comparaison de génomes des polymorphismes spécifiques de la lignée Triticum et de développer des outils de détection, utilisant différentes techniques basées sur l'amplification d'ADN, pouvant discriminer cette lignée à un niveau infraspécifique.

Au sein de la lignée Oryza, une structuration des populations à l'échelle globale en un nombre limité de lignées a pu être mise en évidence par des études antérieures à cette thèse. Toutefois, un parfait consensus concernant cette structure n'est pas atteint entre ces études certainement dû à la variabilité de la taille des échantillons étudiées (nombre d'isolats variant de 45 à 1372), au nombre de marqueurs génétiques inclus (de 10 microsatellites à 751 600 SNP extraits de génomes complets), ainsi qu'aux lieux d'échantillonnages différents des isolats. De plus, les facteurs impliqués dans la mise en place et le maintien de cette structure génétique n'ont été que peu abordés. Les objectifs principaux de cet axe de la thèse étaient de résoudre la structure génétique de cette lignée à l'échelle globale grâce un large échantillonnage (886 isolats prélevés dans toutes les régions rizicoles du monde) et un génotypage de plus de 5.657 SNP répartis dans l'ensemble du génome de *P. oryzae*, d'étudier l'histoire de la colonisation de l'agent pathogène et d'identifier les barrières aux flux de gènes qui sous-tendent cette structure en évaluant : (i) l'allopatrie des lignées ; (ii) les barrières à la reproduction sexuée ; (iii) ainsi que des adaptations différentielles à des facteurs environnementaux que sont la température et la plante hôte.

Ce manuscrit de thèse propose en fin d'introduction un article publié dans la revue Mbio et porté par Pierre Gladieux auquel j'ai participé par la recherche de correspondance entre structure génétique et spectre de virulence. Cet article est à la base du travail proposé dans le chapitre 2 de cette thèse. La suite du manuscrit est structurée en deux chapitres.

Le chapitre 1 porte sur la détection infraspécifique de la lignée Triticum et comporte deux articles. Un article accepté pour publication dans la revue Plant Disease portant sur le développement d'une approche de comparaison de génomes pour la recherche de polymorphismes spécifiques à la lignée Triticum et d'un outil de détection par PCR en temps réel ; et un article en préparation portant sur la mise au point de nouveaux tests de détection par d'autres approches moléculaires, ciblant de nouveaux polymorphismes et permettant la détection sur grains de blé.

Le chapitre 2 de cette thèse porte sur l'étude de la structure génétique de la lignée Oryza et des facteurs impliqués dans cette structuration ainsi que sur de la propagation de l'agent pathogène à l'échelle globale. Ce chapitre se compose d'un article en préparation.

Article 1 : Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen

Les principaux objectifs de cet article étaient de :

- Valider grâce à des données génomiques la structure génétique au sein de la lignée Oryza précédemment observée avec des marqueurs microsatellites ;
- Etablir une phylogénie de ces lignées ;
- Etudier le flux de gène entre ces lignées ;
- Etablir la datation de l'ancêtre commun à l'ensemble de ces lignées ;
- Etudier la correspondance entre structure génétique et spectre de virulence ;

Ma contribution dans cette publication a été de proposer ce dernier axe de recherche et de faire les analyses préliminaires associant génotype et pathotype des isolats.

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

Coexistence of Multiple Endemic and Pandemic Lineages of the Rice Blast Pathogen

Pierre Gladieux,[a] Sébastien Ravel,[a] Adrien Rieux,[b] Sandrine Cros-Arteil,[a] Henri Adreit,[a] Joëlle Milazzo,[a] Maud Thierry,[a] Elisabeth Fournier,[a] Ryohei Terauchi,[c] Didier Tharreau[a]

[a]UMR BGPI, Univ Montpellier, INRA, CIRAD, Montpellier SupAgro, Montpellier, France
[b]CIRAD, UMR PVBMT, St. Pierre de la Reunion, France
[c]Iwate Biotechnology Research Center, Kitakami, Iwate, Japan

ABSTRACT The rice blast fungus *Magnaporthe oryzae* (syn., *Pyricularia oryzae*) is both a threat to global food security and a model for plant pathology. Molecular pathologists need an accurate understanding of the origins and line of descent of *M. oryzae* populations in order to identify the genetic and functional bases of pathogen adaptation and to guide the development of more effective control strategies. We used a whole-genome sequence analysis of samples from different times and places to infer details about the genetic makeup of *M. oryzae* from a global collection of isolates. Analyses of population structure identified six lineages within *M. oryzae*, including two pandemic on japonica and indica rice, respectively, and four lineages with more restricted distributions. Tip-dating calibration indicated that *M. oryzae* lineages separated about a millennium ago, long after the initial domestication of rice. The major lineage endemic to continental Southeast Asia displayed signatures of sexual recombination and evidence of DNA acquisition from multiple lineages. Tests for weak natural selection revealed that the pandemic spread of clonal lineages entailed an evolutionary "cost," in terms of the accumulation of deleterious mutations. Our findings reveal the coexistence of multiple endemic and pandemic lineages with contrasting population and genetic characteristics within a widely distributed pathogen.

IMPORTANCE The rice blast fungus *Magnaporthe oryzae* (syn., *Pyricularia oryzae*) is a textbook example of a rapidly adapting pathogen, and it is responsible for one of the most damaging diseases of rice. Improvements in our understanding of *Magnaporthe oryzae*'s diversity and evolution are required to guide the development of more effective control strategies. We used genome sequencing data for samples from around the world to infer the evolutionary history of *M. oryzae*. We found that *M. oryzae* diversified about 1,000 years ago, separating into six main lineages: two pandemic on japonica and indica rice, respectively, and four with more restricted distributions. We also found that a lineage endemic to continental Southeast Asia displayed signatures of sexual recombination and the acquisition of genetic material from multiple lineages. This work provides a population-level genomic framework for defining molecular markers for the control of rice blast and investigations of the molecular basis of differences in pathogenicity between *M. oryzae* lineages.

KEYWORDS clonality, deleterious mutations, indica rice, introgression, japonica rice, population genomics, population structure, recombination, rice blast, tip-dating calibration

Fungal plant pathogens provide many examples of geographically widespread, often clonal, lineages capable of adapting rapidly to anthropogenic changes, such as the use of new fungicides or resistant varieties, despite extremely low levels of population

Received 3 October 2017 **Accepted** 26 February 2018 **Published** 3 April 2018

Citation Gladieux P, Ravel S, Rieux A, Cros-Arteil S, Adreit H, Milazzo J, Thierry M, Fournier E, Terauchi R, Tharreau D. 2018. Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen. mBio 9:e01806-17. https://doi.org/10.1128/mBio.01806-17.

Editor David Guttman, University of Toronto

Address correspondence to Pierre Gladieux, pierre.gladieux@inra.fr.

genetic diversity (1, 2). An accurate characterization of the population biology and evolutionary history of these organisms is crucial to an understanding of the factors underlying their emergence and spread and to provide new, powerful, and enduring solutions to control these factors. Knowledge of the origins and lines of descent connecting extant pathogen populations provides insight into the pace and mode of disease emergence and subsequent dispersal (2, 3). By inferring the history and structure of pathogen populations, we can also identify disease reservoirs and improve our understanding of the transmissibility and longevity of populations (4, 5). Finally, quantification of the amount and distribution of genetic variation across space and time provides a population-level genomic framework for defining molecular markers for pathogen control and for investigations of the molecular basis of differences in phenotype and fitness between divergent pathogen lineages.

Rice blast is one of the most damaging rice diseases worldwide (6–8). It is caused by the ascomycete fungus *Magnaporthe oryzae* (syn., *Pyricularia oryzae*), which has become a model for plant pathology in parallel with the development of rice as a model crop species (7, 9–11). The rice-infecting lineage of *M. oryzae* coexists with multiple host-specialized and genetically divergent lineages that infect other cereals and grasses (12–14). The lineage infecting foxtail millet (*Setaria italica*, referred to hereafter as *Setaria*) is the closest relative of the rice-infecting lineage, and rice blast was thus thought to have emerged following a host shift from *Setaria* about 2,500 to 7,500 years ago (15), at a time when *Setaria* was the preferred staple in East Asia (16, 17). *Magnaporthe oryzae* infects the two major subspecies of rice, *Oryza sativa* subsp. *indica* and *Oryza sativa* subsp. *japonica* (referred to here as indica and japonica, respectively). Population genomics studies have provided support for a model in which *de novo* domestication occurred only once, to generate the japonica lineage, which subsequently diverged into temperate and tropical japonica, with introgressive hybridization from japonica leading to domesticated indica (18–20). Using microsatellite markers, Saleh et al. (21) identified multiple endemic and pandemic genetic pools of rice-infecting strains, but they were unable to resolve the evolutionary relationships between them. Rice blast has proved able to adapt rapidly to varietal resistance and is thus a dynamic threat to such resistance in rice agrosystems (22). This ability to adapt is surprising given the low level of diversity in *M. oryzae* and its infertility or asexual mode of reproduction in most rice-growing areas (22, 23). This pathogen may thus be particularly exposed to the "cost of pestification" (by analogy with the cost of domestication [24–27]), according to which the combination of a small effective population size, strong selection on pestification genes, and a lack of recombination lead to the accumulation of deleterious mutations (28). Potential limitations to adaptation could be counterbalanced by boom-and-bust cycles in *M. oryzae*, with adaptation occurring during the boom phases, when the short-term effective population size is large (2, 29). Adaptive mutations may also be introduced by cryptic genetic exchanges with conspecifics or heterospecifics (30–33), but these mechanisms remain to be investigated in natural populations of *M. oryzae* (34). An accurate understanding of the population genetics of successful clonal fungal pathogens, such as *M. oryzae*, can provide important insights into the genomic and eco-evolutionary processes underlying pathogen emergence and adaptation to anthropogenic changes.

We used pathogenicity data and whole-genome resequencing data for *M. oryzae* samples distributed over time and space to address the following questions. What population structure does *M. oryzae* display? Does this species consist of relatively ancient or recent clonal lineages? What is the history of temperate japonica, tropical japonica, and indica japonica rice colonization by *M. oryzae*? Do *M. oryzae* lineages display differences in pathogenicity toward rice subspecies? Can we identify genetic exchanges between rice-infecting lineages and the genomic regions that have been exchanged? Is there evidence for a cost of pestification in terms of the accumulation of deleterious mutations?

RESULTS

Genome sequencing and SNP calling. We elucidated the emergence, diversification, and spread of *M. oryzae* in rice agrosystems by studying genome-wide variation across geographically widespread samples. We used 25 and 18 genomes sequenced by Illumina single-end and paired-end read technologies, respectively, with 7 published genome sequences obtained via Solexa and mate-pair titanium methods (10, 12). We thus had a total of 50 genomes available for analysis (see Table S1 in the supplemental material). Forty-five of the isolates concerned originated from cultivated rice (*Oryza sativa*), four from cultivated barley (*Hordeum vulgare*), and one from foxtail millet (*Setaria italica*). The sample set included multiple samples from geographically separated areas (North and South America, South, Southeast, and East Asia, sub-Saharan Africa, Europe, and the Mediterranean), and the reference laboratory strain 70-15 and its parent GY11 were from French Guiana. Nine samples were collected from tropical japonica rice, 7 from temperate japonica, 15 from indica, and 3 from hybrid elite varieties. Sequencing reads were mapped onto the 41.1-Mb reference genome of strain 70-15. Mean sequencing depth ranged from 5× to 64× for genomes sequenced with single-end reads and from 5× to 10× for genomes sequenced with paired-end reads (Table S1). Single-nucleotide polymorphism (SNP) calling identified 182,804 biallelic SNPs distributed over seven chromosomes. The data set consisted of 95,925 SNPs, excluding the *Setaria*-infecting lineage, 61,765 of which had less than 30% missing data and 16,370 of which had no missing data.

Population subdivision, genealogical relationships, and levels of genetic variation. We used a multivariate analysis of population subdivision method, rather than model-based clustering algorithms, because multivariate methods require no assumptions about outcrossing, random mating, or linkage equilibrium within clusters, and previous studies have shown that, in many populations, *M. oryzae* has lost its sexual recombination capacity (references 21 to 23 and references therein). We used a discriminant analysis of principal components (DAPC) to determine the number of lineages represented in our data set. When we progressively increased the number of clusters (*K*) from 2 to 5, we identified the four lineages previously described by Saleh et al. in Asia (21) on the basis of microsatellite data, and we also identified a cluster of three strains collected from the Yunnan and Hunan provinces of China (Fig. 1). Further increases in *K* led to the subdivision of this Yunnan-Hunan cluster. Barley-infecting isolates clustered within rice-infecting lineage 1, which confirmed findings of previous phylogenetic studies (12, 13). Barley is "universally susceptible" to rice-infecting isolates, at least under laboratory conditions. However, the barley isolates included in this study were collected in Thailand, and no major blast epidemic has since been reported on this host in this area, indicating that barley is a minor host for rice-infecting populations.

We investigated whether the clusters observed at *K* values of >4 in the DAPC represented new independent lineages or subdivisions of the main clusters by using RAxML to infer a genome genealogy (35). We based the analysis on a data set combining the full set of SNPs and monomorphic sites, rather than just SNPs, to increase topological and branch length accuracy (36). The total evidence genealogy revealed the existence of four lineages, corresponding to lineages 1 to 4 described by Saleh et al. (21), and two new lineages (named lineages 5 and 6) corresponding to the three-individual cluster observed at $K = 5$ in the DAPC (Fig. 1). With the 41-Mb data set, including missing data, the most basal divergence within the rice-infecting lineage was that between lineage 1 and the other five lineages (Fig. 1). If positions with missing data were excluded (15 Mb), the most basal divergence was that between a group composed of lineages 1, 2, and 6 and a group composed of lineages 3, 4, and 5 (data not shown).

Absolute divergence (d_{xy}) between pairs of lineages was on the order of 10^{-4} differences per base pair and was highest in comparisons with lineage 6 (Table S2). Nucleotide diversity within lineages was an order of magnitude lower than divergence

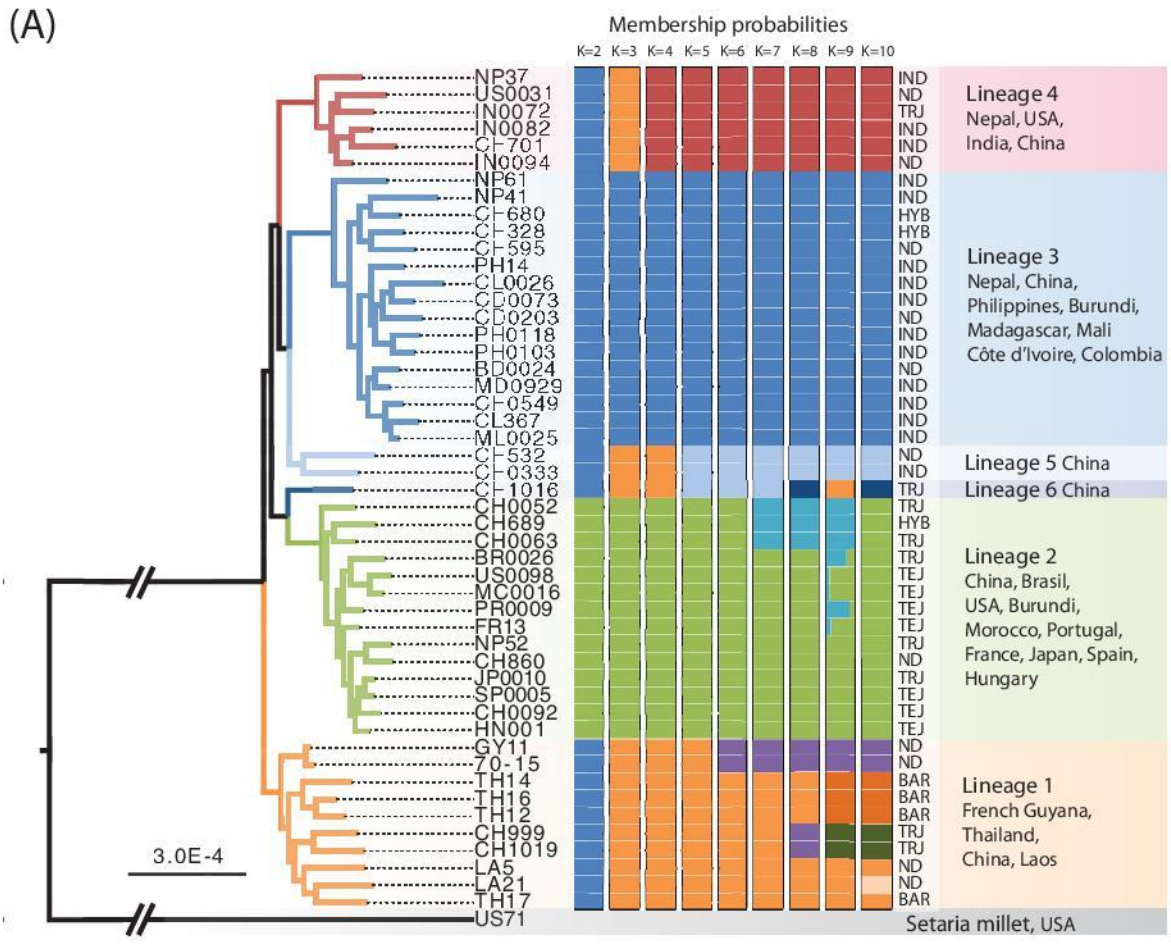


FIG 1 Population subdivision in the sample set analyzed. (A) Total evidence maximum-likelihood genome genealogy and discriminant analysis of principal components. (B) Geographic distribution of the six lineages identified, based on results presented in panel A. In panel A, all nodes had more than 95% bootstrap support (100 resamplings), except for the node carrying isolates BR0026, US0098, PR0009, and MC0016 (support, 72%). On the bar plot, each isolate is represented by a thick horizontal line divided into *K* segments, indicating the isolate's estimated probability of belonging to the *K* assumed clusters. In panel B, diameters are proportional to the number of isolates collected per site (the smallest diameter represents 1 isolate). TRJ, tropical japonica; TEJ, temperate japonica; IND, indica; HYB, hybrid; BAR, barley; ND, no data.

TABLE 1 Summary of population genomic variations in nonoverlapping 100-kb windows[a]

Lineage	*n*	*S*	*K*	H_e	θ_w	π	D
1	10	57.6	3.6	0.31	2.25E−04	2.11E−04	−0.558
2	14	19.7	7.2	0.20	6.94E−05	4.92E−05	−1.454
3	16	21.5	7.9	0.17	7.24E−05	4.53E−05	−1.718
4	6	10.5	4.3	0.38	5.15E−05	4.52E−05	−0.824

[a]Lineages 5 and 6 were not included in calculations because the sample sizes for these lineages were too small ($n = 2$ and $n = 1$, respectively). *n*, sample size; θ_w, Watterson's θ per base pair; π, nucleotide diversity per base pair; H_e, haplotype diversity; *K*, number of haplotypes; D, Tajima's neutrality statistic.

in lineages 2 to 4 (θ_w per site, 5.2e−5 to 7.2e−5; π per site, 4.5e−5 to 4.9e−5) and was highest in lineage 1 (θ_w per site, 2.3e−4; π per site, 2.1e−4) (Table 1). Tajima's D was negative in all lineages, indicating an excess of low-frequency polymorphisms, and values were closer to zero in lineages 1 and 4 (D = −0.56 and −0.82, respectively) than in lineages 2 and 3 (D = −1.45 and −1.72, respectively). The same differences in levels of variability across lineages, and individual summary statistics of the same order of magnitude, were observed if missing data were excluded from computations.

Footprints of natural selection and the cost of pestification. We tested for standard neutral molecular evolution by using the McDonald-Kreitman method, based on genome-wide patterns of synonymous and nonsynonymous variations (Table 2). The null hypothesis could be rejected for all four lineages ($P < 0.0001$). The neutrality index, which quantifies the direction and degree of departure from neutrality, was greater than 1, indicating an excess of amino acid polymorphisms. This pattern suggests that lineages 1 to 4 accumulated slightly deleterious mutations during divergence from the *Setaria*-infecting lineage. Under near-neutrality, the ratio of nonsynonymous to synonymous nucleotide diversity (π_N/π_S) provides an estimate of the proportion of effectively neutral mutations that are strongly dependent on the effective population size, N_e (37). The π_N/π_S ratio ranged from 0.43 in lineage 1 to 0.61 in lineage 4 and was intermediate in lineages 2 and 3 ($\pi_N/\pi_S = 0.49$), and the ratio of nonsense (i.e., premature stop codons) to sense nonsynonymous mutations ($P_{nonsense}/P_{sense}$) followed the same pattern. Overall, the π_N/π_S and $P_{nonsense}/P_{sense}$ ratios obtained suggest a higher proportion of slightly deleterious mutations segregating in lineage 4 and, to a lesser extent, in lineages 2 and 3, than in lineage 1. Assuming identical mutation rates, we can estimate that the long-term population size of lineage 1 ($\pi_S = 0.00018$/bp) was 2.5 to 3 times greater than that of the other lineages, consistent with the effect of N_e on the efficacy of negative selection predicted under near-neutrality.

Distribution and reproductive biology of *M. oryzae* lineages. The strains of lineages 1 and 2 originated from rain-fed upland rice, including rice grown in experimental fields. Lineage 2 was exclusively associated with tropical and temperate japonica, whereas lineage 1 was sampled from barley, tropical japonica, and hybrid rice varieties (Fig. 1; Table S1). Lineage 1 was restricted to continental Southeast Asia (Laos, Thailand, Yunnan). The reference laboratory strain GY-11 (also referred to as Guy11) was

TABLE 2 Results of McDonald-Kreitman tests based on genome-wide patterns of synonymous and nonsynonymous variation and measurements of the genome-wide intensity of purifying selection[a]

Lineage	π_N/π_S	$P_{nonsense}/P_{sense}$	P_n/P_s	D_n/D_s	NI
1	0.43 (0.00041/0.00018)	0.011 (49/4,244)	1.23 (4,293/3,492)	0.70 (16,444/23,656)	1.77*
2	0.49 (0.00012/0.00006)	0.022 (36/1,622)	1.52 (1,658/1,088)	0.72 (15,565/21,745)	2.13*
3	0.49 (0.00015/0.00007)	0.018 (32/1814)	1.17 (1,846/1,578)	0.97 (14,789/15,293)	1.21*
4	0.61 (0.00012/0.00007)	0.034 (31/914)	1.59 (945/593)	0.72 (15,302/21,347)	2.22*

[a]Divergence was measured against predicted gene sequences of the *Setaria*-infecting *Magnaporthe oryzae* isolate US71. π_N/π_S is the ratio of nonsynonymous to synonymous nucleotide diversity. Under near-neutrality, π_N/π_S provides an estimate of the proportion of effectively neutral mutations strongly dependent on effective population size, N_e. π_S is a proxy for N_e. $P_{nonsense}/P_{sense}$ is the number of nonsynonymous nonsense mutations (e.g., a "premature" stop codon) divided by the number of nonsynonymous sense mutations. The neutrality index (NI) = $(P_n/P_s)/(D_n/D_s)$ and determines the direction and degree of departure from neutrality; *, $P < 0.0001$, chi-square test of independence. NI is equal to 1 if nonsynonymous mutations are neutral or strongly deleterious. NI is <1 when amino acid substitutions have occurred and implies that advantageous mutations have become fixed. NI is >1 when there is an excess of amino acid polymorphisms, as expected in a context of slightly deleterious mutations.

collected in French Guiana, from fields cultivated by Hmong refugees who fled Laos in the 1970s. Lineage 2 was pandemic and included all the European samples.

Lineage 3 and 4 samples originated from irrigated or rain-fed upland/lowland rice. They were mostly associated with indica rice, with two samples collected from hybrid varieties and one collected from tropical japonica (Fig. 1; Table S1). Lineage 3 was pandemic and was found in all sub-Saharan Africa samples, whereas lineage 4 was found on the Indian subcontinent, in Zhejiang (China), and the United States. Lineages 5 and 6 were collected from indica and tropical japonica varieties of rain-fed upland rice in Yunnan and Hunan, China, respectively.

Lineages 2, 3, and 4 displayed low rates of female fertility (20%, 0%, and 0%, respectively) and a significant imbalance in mating type ratio (frequency of Mat-1, 100%, 14.3%, and 100%, respectively; chi-square test, $P < 0.001$), whereas lineage 1 had a female fertility rate of 88.9% and a nonsignificant imbalance in mating type ratio (frequency of Mat-1, 33.3%; chi-square test, $P = 0.083$). Lineage 5 was Mat-1, and only one of the two strains was female fertile (no data for lineage 6).

Pathogen compatibility range. Gallet et al. (38) analyzed the range of compatibility, in terms of the qualitative success of infection, between 31 *M. oryzae* isolates and 57 rice genotypes. Analyses of variance revealed a pattern of host-pathogen compatibility strongly structured by the host of origin of the isolates (i.e., the rice subspecies from which samples were collected). We investigated whether the compatibility between rice hosts and *M. oryzae* isolates was also structured by the lineage of origin of the isolates, by supplementing the data set published by Gallet et al. (38) with pathotyping data for 27 isolates. We added microsatellite data to the SNP data, to overcome the absence of sequence data for 28 isolates, and we used clustering methods to confidently assign 46 of the 58 isolates with pathotyping data to identified lineages (no isolates could be assigned to lineage 5 or 6 [see Materials and Methods]). The final pathogenicity data set included 46 isolates from lineages 1 to 4, inoculated onto 38 tropical japonica, temperate japonica, and indica varieties and 19 differential varieties with known resistance genes (Table S3).

Infection success (binary response) was analyzed with a generalized linear model. An analysis of the proportion of compatible interactions revealed significant effects of rice subspecies, pathogen lineage, and the interaction between them (Table S4). The lineage effect could be explained by lineage 2 having a lower infection frequency than lineage 1 (comparison of lineages 1 and 2: $z = -2.779$, $P = 0.005$) and by lineage 3 having a higher infection frequency than lineage 1 (comparison of lineages 3 and 1: $z = 2.683$, $P = 0.007$), whereas the infection frequency of lineage 4 was not significantly different from that of lineage 1 (comparison of lineages 4 and 1: $z = 1.121$, $P = 0.262$). The rice subspecies effect could be attributed to tropical japonica varieties having a wider compatibility range than indica varieties (comparison of tropical japonica and indica: $z = 1.793$, $P = 0.073$) and temperate japonica having a wider compatibility range than indica varieties (comparison of temperate japonica and indica: $z = 1.830$, $P = 0.067$). The significant interaction between rice subspecies and pathogen lineage indicates that the effect of the lineage of origin of the isolate on the proportion of compatible interactions differed between the three rice subspecies. This interaction effect can be attributed to pathogen specialization on indica and tropical japonica, with lineage 1 (mostly originating from tropical japonica or from areas in which tropical japonica is grown) infecting tropical japonica varieties more frequently than indica varieties, lineage 2 (the lineage sampled from temperate japonica) infecting temperate japonica varieties more frequently than other varieties, lineages 3 and 4 (mostly originating from indica varieties) infecting indica varieties more frequently than tropical japonica varieties, and all four lineages infecting temperate japonica varieties at relatively high frequencies (Fig. 2A; Table S4).

Major resistance (R) genes can be a major determinant of pathogen host range, and they promote divergence between pathogen lineages by exerting strong divergent selection on a limited number of pathogenicity-related genes (39–41). We investigated

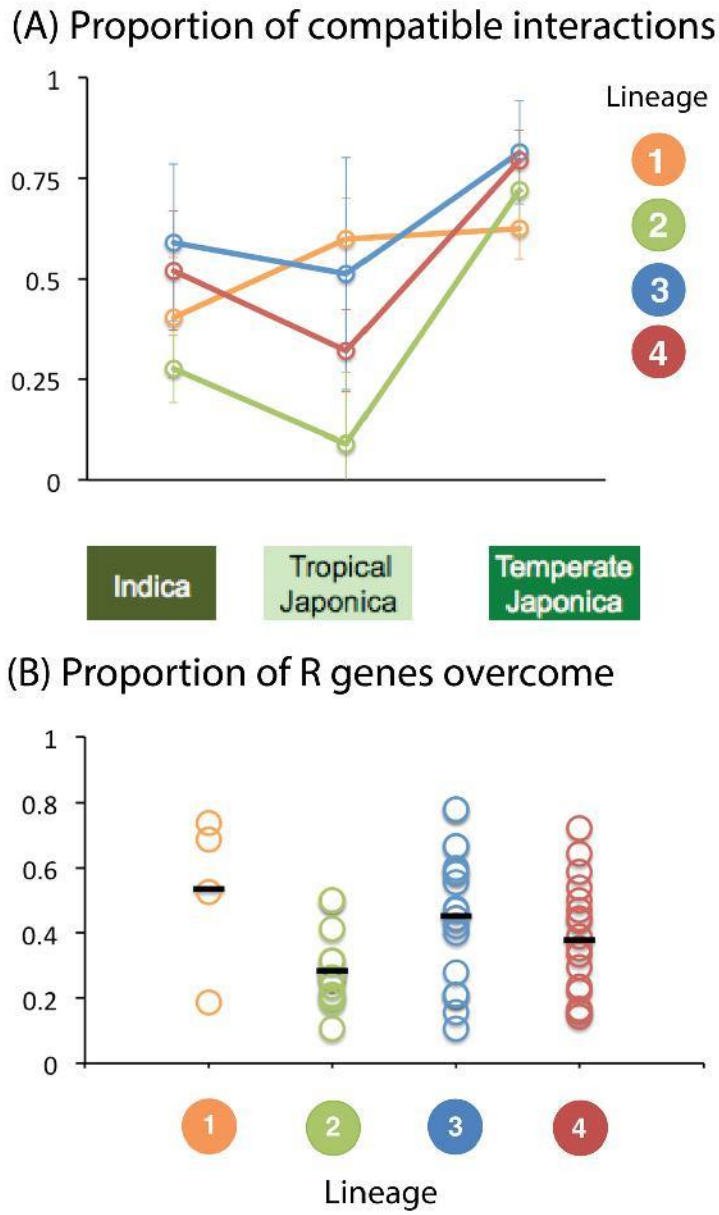


FIG 2 Proportion of compatible interactions between 46 isolates from lineages 1 to 4 of *M. oryzae* and 38 varieties representing three rice subspecies (A) and the proportion of R genes overcome by 36 isolates from lineages 1 to 4 of *M. oryzae* used to inoculate 19 differential lines of rice (B).

the possible role of major resistance genes in the observed differences for compatibility between rice subspecies and pathogen lineages by challenging 19 differential varieties with the 46 isolates assigned to lineages 1 to 4. An analysis of the number of R genes overcome revealed a significant effect of pathogen lineage (Table S5). This effect was driven mostly by lineage 2, which overcame fewer R genes than the other lineages (Fig. 2B; Table S5).

Recombination within and between lineages. We visualized evolutionary relationships while taking into account the possibility of recombination within or between lineages by using the phylogenetic network approach Neighbor-Net, as implemented in Splitstree 4.13 (42). Neighbor-Net is an agglomerative method that generates planate split graph representations. A split is a partitioning of the data set, and a collection of splits is considered compatible if they fall within the set of splits of a tree. Gene genealogies represent compatible collections of splits, whereas Neighbor-Net can be

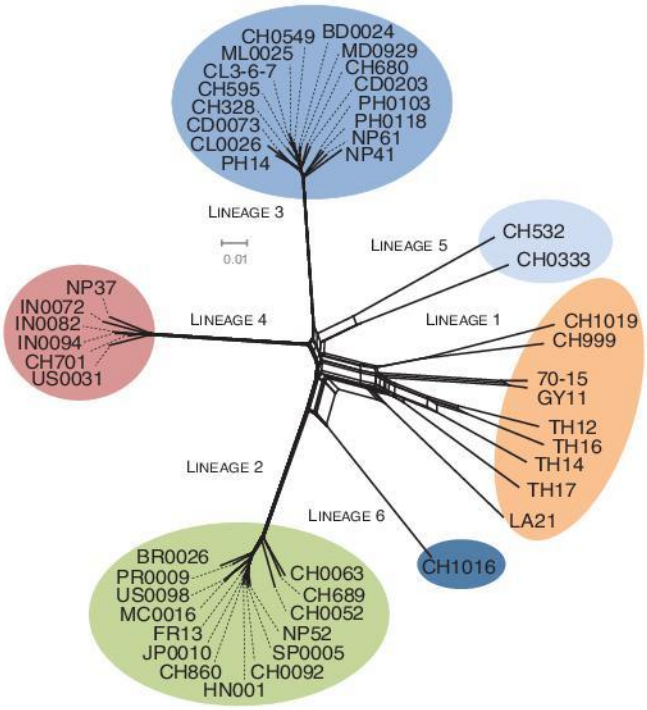


FIG 3 Neighbor-Net networks showing relationships between haplotypes identified on the basis of the full set of 16,370 SNPs without missing data in the whole sample set (A), in lineage 1 (B), in lineage 2 (C), in lineage 3 (D), and in lineage 4 (E).

used to visualize conflicting phylogenetic signals, represented by network reticulation, through a condition weaker than compatibility. The Neighbor-Net network inferred from the set of 16,370 SNPs without missing data presented a non-tree-like structure of the inner connections between lineages, consistent with genetic exchanges between unrelated isolates or incomplete lineage sorting (Fig. 3). Greater network reticulation was observed between lineages 1, 5, or 6 and the other lineages than between these other lineages themselves. Lineages 2 to 4 had long interior branches and star-like topologies, consistent with long-term clonality.

We evaluated the amount of recombination within lineages by estimating the population recombination parameter (rho = $2 N_e r$) and testing for the presence of recombination with a likelihood permutation test implemented in the Pairwise program in LDHat. Recombination analyses confirmed the heterogeneity between lineages of the contribution of recombination to genomic variation, with recombination rates averaged across chromosomes of more than 2 to 3 orders of magnitude higher in lineage 1 (10.57 crossovers/Mbp/generation) than in other lineages (lineage 2, 0.28; lineage 3, 0.01; lineage 4, 0.33 crossovers/Mbp/generation) (Table 3). SplitsTree analy-

TABLE 3 Estimates of the population recombination rate (ρ), tests of recombination based on homoplasy and linkage disequilibrium, and the proportion of homoplastic SNPs

Lineage	ρ (no. of crossovers/Mbp/generation) on chromosome[a]								% homoplastic SNPs	phi test *P* value
	1	2	3	4	5	6	7	Mean		
1	8.6*	3.8*	15.1*	1.4	8.5*	10.6*	13.5*	10.57	34.56	0.0000
2	0.0	0.2	0.2*	0.3	0.6	0.5	0.3	0.28	0.09	0.0944
3	0.4*	0.2*	0.0	0.0	0.0	0.0	0.0	0.01	0.47	0.0535
4	0.2	0.2*	0.3	0.4	0.4	0.4	0.4	0.33	0.40	0.0014

[a]*, $P < 0.05$. The phi test assesses pairwise homoplasy. The null hypothesis of no recombination was tested, with the phi test and for ρ, using random permutations of the positions of the SNPs based on the expectation that sites are exchangeable if there is no recombination. For the ρ test, significance was determined from the distribution of maximum composite likelihood values calculated from permuted data.

ses, producing the reticulations within each lineage and testing for recombination with the phi test, were consistent with this pattern (Table 3 and Fig. 3; Fig. S1). The null hypothesis of no recombination was rejected only for lineages 1 and 4 (43).

Differences in recombination-based variation between lineages were confirmed by analyses of homoplasy (Table 3). Homoplastic sites display sequence similarities that are not inherited from a common ancestor; instead, they result from independent events in different branches. Homoplasy can result from recurrent mutations or recombination, and the contribution of recombination to homoplasy is expected to predominate in outbreeding populations. Homoplastic sites were identified by mapping mutations onto the total evidence genome genealogy with the Trace All Characters function of Mesquite (44), applying ancestral reconstruction under the maximum parsimony optimality criterion. The resulting matrix of ancestral states for all nodes was then processed with a python script to determine the number of mutations that had occurred at each site within each lineage, counting sites displaying multiple substitutions across the tree as homoplastic. Only 0.09%, 0.47%, and 0.40% of the SNPs were homoplastic in lineages 2, 3, and 4, respectively, versus 34.6% in lineage 1 (lineages 5 and 6 were not tested due to the small sample sizes). The very small numbers of homoplastic sites in lineages 2, 3, and 4 suggested that these lineages are largely clonal, whereas the high level of homoplasy detected in lineage 1 is consistent with repeated recombination events between strains of this lineage.

We assessed the genomic impact of recombination by analyzing patterns of linkage disequilibrium (LD), i.e., the tendency of different alleles to occur together in a nonrandom manner. For lineage 1 ($S = 13{,}000$ SNPs), LD decayed smoothly with physical distance, reaching half its maximum value at about 10 kb, whereas for lineages 2, 3, and 4 ($S = 3{,}700$, 3,200, and 2,700 SNPs), no LD decay pattern was observed (Fig. S2). These analyses also revealed that background LD levels were no higher in lineages 2, 3, or 4, which appeared to be largely clonal, than in lineage 1. However, both simulation work and empirical data have shown that population history, including bottlenecks and admixtures, strongly affects the background level of LD in a population (45).

Genome scan for genetic exchanges between lineages. We scanned the genomes for the exchange of mutations between lineages, using a method based on lineage-diagnostic SNPs and a probabilistic method of "chromosome painting" (Fig. 4). In the lineage-diagnostic SNP approach, each isolate is removed from the data set in turn to identify SNPs specific to a particular lineage (i.e., biallelic sites displaying a mutation specific to a given lineage). Each focal isolate is then added back to the data set and scanned for the presence of lineage-diagnostic SNPs identified in lineages other than its lineage of origin. Using this approach, we identified 515 lineage-diagnostic singletons with 276, 96, and 140 singletons in lineages 1, 5, and 6, respectively, and only 1 singleton in each of lineages 2, 3, and 4. Putatively migrant singletons were assigned to all other lineages for lineages 1 and 5 and to all other lineages except lineage 1 for lineage 6 (Table S6). Chromosome painting is a probabilistic method for reconstructing the chromosomes of each individual sample as a combination of all other homologous sequences. We identified the migrant mutations present in each isolate, with these mutations being defined as those having a probability greater than 90% of resulting from being copied from a lineage other than the lineage of origin of the focal isolate. This method uses population data from recipient populations only, and we were therefore able to include only lineages 1 to 4 in the analysis. Chromosome painting identified 464 migrant mutations, all of which segregated in lineage 1. Putative migrant mutations were assigned to all five of the other lineages (92.8 mutations per lineage, on average), with lineage 2 making the largest contribution (165 mutations) and lineage 4 the smallest contribution (39 mutations).

The sets of putative migrant mutations identified by the two methods matched different sets of genes enriched in NOD-like receptor (NLR) (46), HET domain (47), or the GO term lipid catabolic process (Table S6). However, the presence of false positives due to the random sorting of ancestral polymorphisms in lineage 1 and other lineages

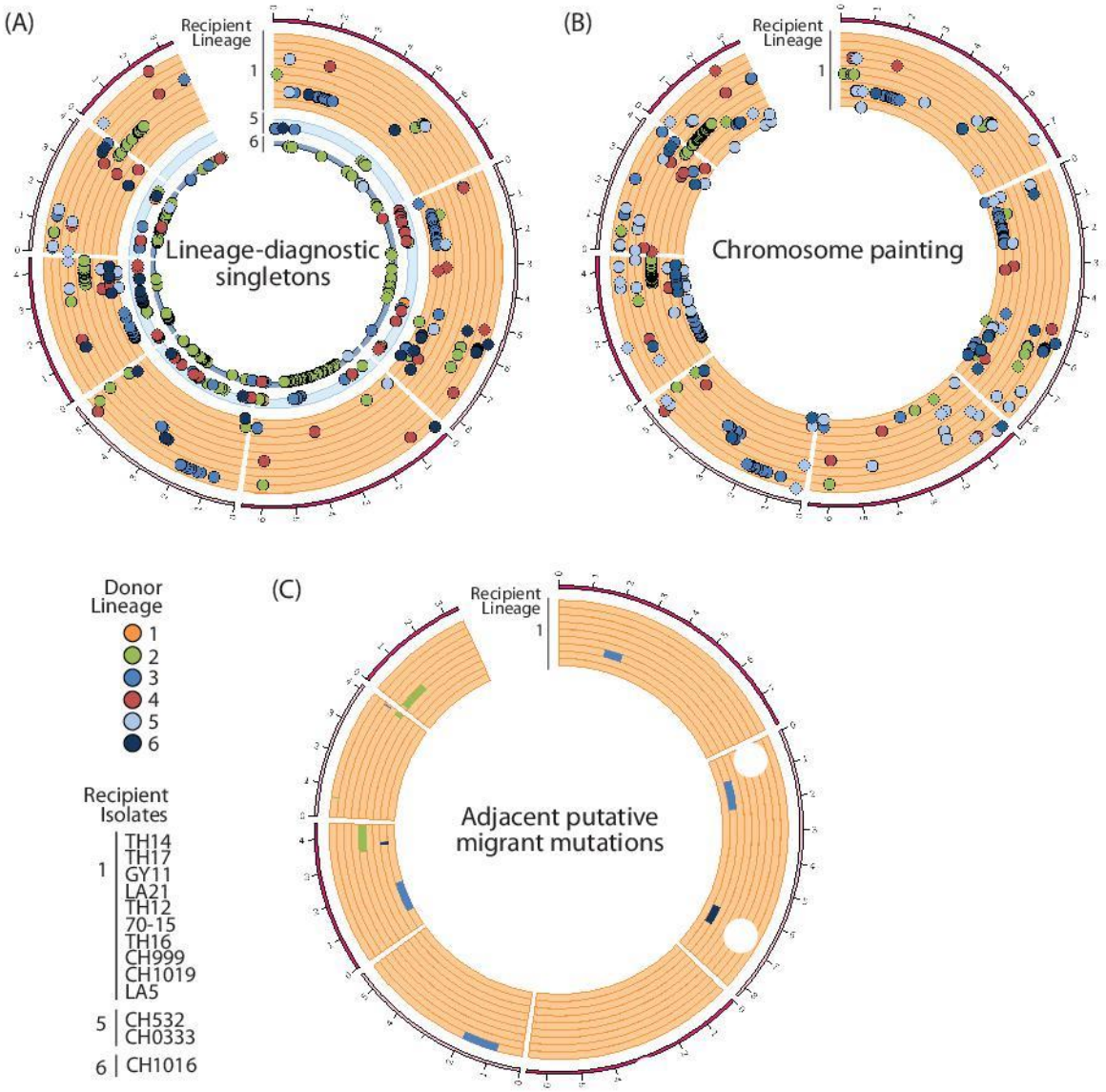


FIG 4 Genomic distribution of candidate immigrant mutations in lineages 1, 5, and 6. (A) Lineage-diagnostic mutations segregating as singletons in other lineages. (B) Lineage 1 mutations for which the most probable donor was lineage 2, 3, or 4 in probabilistic chromosome painting analysis. Lineages 5 and 6 ($n = 2$ and $n = 1$, respectively) could not be included as recipient populations in the chromosome painting analysis due to their small sample sizes. No candidate immigrant mutations were identified in lineages 2, 3, or 4. (C) Genomic regions corresponding to a series of adjacent putative migrant mutations identified with lineage-diagnostic singletons in lineage 1. Chromosomes 1 to 7 appear in clockwise order, with ticks at megabase intervals.

cannot be excluded. We minimized the impact of the retention of ancestral mutations by reasoning that series of adjacent mutations are more likely to represent genuine gene exchange events. We identified all the genomic regions defined by three adjacent putative migrant mutations originating from the same donor lineage. We searched for such mutations among the set of putative migrant mutations identified by the two methods. We identified 12 such regions in total, corresponding to 1,917 genes. Functional enrichment tests for each recipient isolate revealed enrichment of genes for the GO term pathogenesis for isolate CH999, the GO term phosphatidylinositol biosynthetic process for isolate TH17, and the GO term telomere maintenance for isolate CH1019 (Table S6).

Molecular dating. We investigated the timing of rice blast emergence and diversification by performing Bayesian phylogenetic analyses with Beast. Isolates were collected from 1967 to 2009 (Table S1), making it possible to use a tip-based calibration

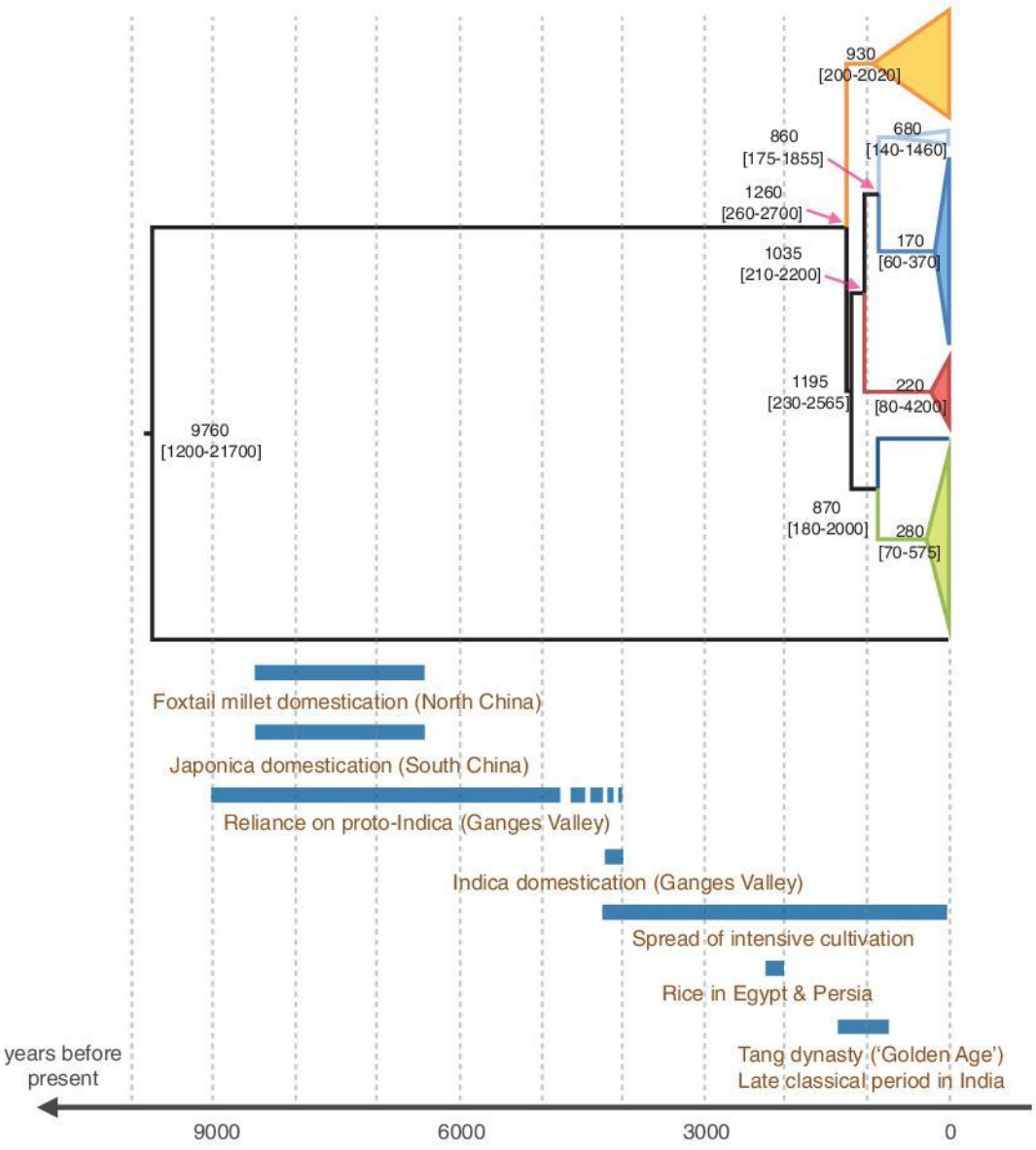


FIG 5 Tip-calibrated genealogy inferred by maximum-likelihood phylogenetic inference using Beast 1.8.2, based on single-nucleotide variations in 50 *M. oryzae* genomes. Approximate historical periods are shown for context.

approach to estimate evolutionary rates and ancestral divergence times together. We analyzed the linear regression of sample age against root-to-tip distance (i.e., the number of substitutions separating each sample from the hypothetical ancestor at the root of the tree). The temporal signal obtained in this analysis was strong enough for thorough tip-dating inferences (Fig. S3) (48). We therefore used tip dating to estimate the rate at which mutations accumulate (i.e., the substitution rate) and the age of every node in the tree, including the root (i.e., time to the most recent common ancestor), simultaneously. At the scale of the genome, the mean substitution rate was estimated at 1.98e−8 substitutions/site/year (Fig. S3). The six rice-infecting lineages were estimated to have diversified ~900 to ~1,300 years ago (95% highest posterior density [HPD], 175 to 2,700 years ago) (Fig. 5). Bootstrap node support was strong, and similar node age estimates were obtained when the recombining lineage 1 and the potentially recombining lineages 5 and 6 (data not shown) were excluded, indicating the limited effect of recombination on our inferences. We also inferred that the ancestor of rice-infecting and *Setaria*-infecting lineages lived ~9,800 years ago. However, the

credibility intervals were relatively large (95% HPD, 1,200 to 22,000 years ago), covering the period from japonica rice domestication and *Setaria* domestication to the last glacial maximum and overlapping with previous estimates suggesting that the rice- and *Setaria*-infecting lineages diverged shortly after rice domestication, or even during the period of rice domestication (range of point estimates in reference 15, 2,500 to 7,300 years ago).

DISCUSSION

We performed a whole-genome sequence analysis of 50 isolates with different temporal and spatial distributions in order to elucidate the emergence, diversification, and spread of *M. oryzae* as a rapidly evolving pathogen with a devastating impact on rice agrosystems. Analyses of population subdivision confirmed the four lineages previously identified by Saleh et al. (21). Previous analyses of microsatellite data were unable to resolve the genealogical relationships between clusters or to capture the phylogenetic depth of population subdivision within *M. oryzae*. In contrast, our population genomic analyses of resequencing data revealed weak divergence between clusters (absolute divergence [d_{xy}] on the order of 10^{-4} differences per base pair), consistent with recent diversification. Phylogenetic analyses using sampling dates for calibration confirmed the recent origin of the six lineages, with estimates of divergence times ranging from ~900 to ~1,300 years ago (95% credible intervals, 175 to 2,700 years ago). Lineage 1 (which includes the reference strains GY11 and 70-15) was found in mainland Southeast Asia and originates from barley, tropical japonica, or undetermined varieties. All isolates from lineages 1, 5, and 6 were collected in rain-fed upland agrosystems typical of japonica rice cultivation, and pathogenicity test results were consistent with the local adaptation of lineage 1 to tropical japonica rice. Lineage 2 was pandemic in irrigated fields of temperate japonica rice outside Asia, and cross-inoculation experiments revealed specialization on this host and an ability to overcome fewer R genes, on average, than other lineages. Lineages 3 and 4 were associated with indica. Lineage 3 is pandemic, and cross-inoculation indicated local adaptation to this host, relative to tropical japonica, although lineages 3 and 4 had relatively wide compatibility ranges, consistent with generalism. One possible explanation for the wide compatibility range of temperate japonica varieties and the narrow compatibility range of lineage 2 is that temperate japonica varieties have smaller repertoires of R genes, as resistance to blast is of less concern to breeders growing rice under temperate irrigated conditions, which are less conducive to epidemics (38).

The continental Southeast Asian lineage was the most basal in total evidence genome genealogies, reflecting a pathway of domesticated Asian rice evolution (16, 18) in which the *de novo* domestication of rice occurred only once, in japonica. However, the diversification of *M. oryzae* into multiple rice-infecting lineages (point estimates ranging from ~900 to ~1,300 years ago) appears to be much more recent than the *de novo* domestication of rice (8,500 to 6,500 years ago [16, 49, 50]), the spread of rice cultivation in paddy fields, and the domestication of indica in South Asia, following introgressive hybridization from the early japonica gene pool into "proto-indica" rice (about 4,000 years ago [16, 51]). At the time corresponding to the upper bound of the 95% credible interval (2,700 years ago), japonica rice and paddy field cultivation had spread to most areas of continental and insular South, East, and Southeast Asia, and indica rice was beginning to spread out of the Ganges plains (16, 52). The point estimates for the splitting of *M. oryzae* lineages correspond to the Tang Dynasty ("the Golden Age") in China and the late classical period in India, during which food production became more rational and scientific and intensive irrigated systems of cultivation were developed, bringing about economic, demographic, and material growth (53).

Genome scans based on polymorphism and divergence revealed heterogeneity in the genomic and life history changes associated with the emergence and spread of the different lineages. Using microsatellite data and a larger collection of samples, Saleh et al. (21) identified differences in variability levels between lineages, with similar or higher

levels of genetic variability in lineages 1 and 4 than in lineages 2 and 3. Lineages 1 and 4 were also the only lineages that displayed biological features (fertile female rates and mating type ratios) consistent with sexual reproduction. Our genome-wide analyses of variability and linkage disequilibrium provided clear evidence that the continental Southeast Asian lineage 1 displays recombination and is genetically diverse, suggesting that sexual reproduction occurs and that long-term population size is relatively high, whereas pandemic lineages 2 and 3 are largely clonal and genetically depauperate, suggesting a lack of sexual reproduction and demographic bottlenecks associated with their emergence in agrosystems. However, population genomic analyses did not confirm the previously reported high variability and capacity for sexual recombination of the South Asia–United States lineage 4 (21), possibly due to differences in sample sizes between studies. The null hypothesis of clonality was not rejected by phi tests for recombination, but both total (θ_w) and average (π) nucleotide diversity, and also the population recombination rate (ρ), were on the same order of magnitude in lineage 4 as in lineages 2 and 3, consistent with a lack of recombination and a small effective population size.

The patterns of polymorphism and diversity at nonsynonymous and synonymous sites indicated that deleterious mutations were particularly abundant in clonal lineages 2 to 4 of *M. oryzae*, with the smaller long-term population size, consistent with a higher cost of pestification in these lineages. The introgression of genetic elements from clonal lineages harboring greater loads of deleterious mutations may counteract the efficient purging of deleterious mutations in the recombining lineage 1 from mainland Southeast Asia and lead to smaller differences in the proportion of nonsynonymous mutations between recombining and clonal lineages. However, the extensive variabilities in the origin and genomic distribution of the detected putative migrant mutations suggest that most of these mutations are false positives, with only a series of adjacent mutations of this type originating from the same donor lineage corresponding to genuine genetic exchange events. Field-scale studies in areas in which different lineages coexist should provide more detailed insights into the relative importance of interlineage recombination and make it possible to determine whether genetic exchanges are driven by positive selection or are an incidental by-product of the sympatric coexistence of interfertile lineages. We hypothesize that the accumulation of deleterious mutations in pandemic clonal complexes and gene flow into sexual lineages during disease emergence and spread are widespread phenomena, which are not due to idiosyncrasies of *M. oryzae*, and we expect these patterns to hold true in other invasive fungal plant pathogens.

An examination of additional isolates from undersampled geographic regions (including Africa and South America), based on sequencing approaches and sampling schemes tailored to detect adaptation from *de novo* mutations, will be required to enhance our understanding of the biogeography of *M. oryzae* and the genetic basis of adaptation in the different *M. oryzae* lineages. Nevertheless, the catalog of variants detected in our study provides a solid foundation for future research into the population genomics of adaptation in *M. oryzae*. Our work also provides a population-level genomic framework for defining molecular markers for the control of rice blast and investigations of the molecular basis of the differences in phenotype and fitness between divergent lineages.

MATERIALS AND METHODS

Genome sequencing and SNP calling. Sequencing libraries were prepared and Illumina HiSeq 2500 sequencing was performed either at Beckman Coulter Genomics (BCG; Danvers, MA, USA) or at the Iwate Biotechnological Research Center (Table S1). Genomic DNA for sequencing at BCG was isolated from 100 mg of fresh mycelium grown in liquid medium. The mycelium was treated with enzymes degrading the cell walls (mainly beta-glucanase) and then incubated in lysis buffer (Triton 2×–1% SDS–100 mM NaCl−10 mM Tris-HCl−1 mM EDTA). Nucleic acids were extracted by treatment with chloroform:isoamyl alcohol (24:1), followed by precipitation overnight in isopropanol. They were then rinsed in 70% ethanol. The nucleic acid extract was treated with RNase A (0.2 mg/ml, final concentration) to remove RNA. The DNA was purified by another round of chloroform:isoamyl alcohol (24:1) treatment. Genomic DNA for sequencing at IBRC was isolated with a protocol adapted from the animal tissue (mouse tail) protocol

available for the Promega Wizard genomic DNA purification kit. Nucleic acids were extracted from 20 mg of fresh mycelium grown in liquid medium, which was ground into powder in liquid nitrogen with a prechilled pestle and mortar. The centrifugation time specified for the mouse tail protocol was increased to 15 min, and centrifugation was carried out at 4°C, after precipitation for 3 h at −20°C. Nucleic acids were resuspended in water, treated with RNase A (0.2 mg/ml, final concentration), purified by treatment with chloroform:isoamyl alcohol (24:1), precipitated overnight in isopropanol supplemented with 0.1 volume of sodium acetate (3 M; pH 5), and rinsed in 70% ethanol.

Sequencing reads were either paired-end reads (read length, 100 nucleotides; insert size, ~500 bp; DNAs sequenced at IBI) or single-end reads (read length, 100 nucleotides; DNAs sequenced by BCG). Reads were trimmed to remove barcodes and adapters and were then filtered to eliminate sequences containing ambiguous base calls. Reads were mapped against the 70-15 reference genome, version 8 (10), with BWA (54) (subcommand alb, option -n 5; subcommand sampe option -a 500). Alignments were sorted with samtools (55), and reads with a mapping quality below 30 were removed. Duplicates were removed with Picard (http://broadinstitute.github.io/picard/). We used Realigner-Targetcreator, Target-creator, and Indelrealigner within the genome analyses toolkit (GATK) (56) to define intervals to target for local realignment and for the local realignment of reads around indels, respectively, and Unified Genotyper to call SNPs. We used GATK's SelectVariants to apply hard filters and to select high-confidence SNPs based on annotation values. Numbers of reference and alternative alleles were calculated with JEXL expressions based on the vc.getGenotype().getAD() command. Variants were selected based on the following parameters: counts of all reads with a MAPQ of 0 below 3.0 (MQ0 in GATK), number of reference alleles + number of alternative alleles $\geq$ 15.0, and number of reference alleles/number of alternative alleles $\leq$ 0.1. With these parameters, SNP calls are limited to positions with relatively high sequencing depths and limited discordance across high-quality sequencing reads. We used a second SNP caller, Freebayes v0.9.10-3-g47a713e (57), to assess the impact of the SNP calling method on the sets of SNPs detected, given the presence in our data set of isolates sequenced at relatively low depth ($<10\times$). We set the -min-alternate-count option to one in Freebayes. When the sample-by-sample Freebayes SNP calls were compared with the GATK SNP calls, after filtration, Freebayes identified 1.63× (standard deviation [SD], 0.28) more SNPs per sample on average than via analyses with GATK, and 92.3% (SD, 2.3) of the SNPs identified with GATK were also identified with Freebayes. The size of the intersection between the sets of SNPs identified by the two methods was negatively correlated with sequencing depth (i.e., the concordance between SNP callers was higher for isolates sequenced less deeply), indicating a minimal impact of isolates sequenced at lower depth on confidence in SNP calls. When the multisample Freebayes SNP calls were compared with the GATK SNP calls, after filtration, 83% of the SNPs identified with GATK were confirmed with Freebayes, and the GATK SNPs that were not confirmed with Freebayes were identified in sets of isolates with a genome-wide sequencing depth of 47.8× on average (SD, 8.2), consistent with a minimal impact of isolates sequenced at lower depth on confidence in SNP calls. High-confidence SNPs were annotated with SnpEff v4.3 (58).

Mating type and female fertility assays. Mating type and female fertility for our lineages had previously been determined (23) or we determined them as previously described (59).

Genealogical relationships and population subdivision. Total evidence genealogy was inferred with RAxML from pseudoassembled genomic sequences (i.e., tables of SNPs converted into a fasta file, using the reference sequence as a template), assuming a general time-reversible model of nucleotide substitution with the Γ model of rate heterogeneity. Bootstrap confidence levels were determined with 100 replicates. DAPC was performed with the Adegenet package in R (60). Sites with missing data were excluded. We retained the first 20 principal components and the first six discriminant functions.

Diversity and divergence. Polymorphism and divergence statistics were calculated with Egglib 3.0.0b10 (61), excluding sites with >30% missing data. The neutrality index was calculated as $(P_n/P_s)/(D_n/D_s)$, where P_n and P_s are the numbers of nonsynonymous and synonymous polymorphisms, and D_n and D_s are the numbers of nonsynonymous and synonymous substitutions, respectively. D_n and D_s were calculated with Gestimator (62) using the *Setaria*-infecting lineage as an outgroup. P_n and P_s were calculated with Egglib.

Linkage disequilibrium and recombination. The coefficient of linkage disequilibrium (r^2) (63) was calculated with Vcftools (64), excluding missing data and sites with minor allele frequencies below 10%. For all lineages, we calculated r^2 values for all pairs of SNPs less than 100 kb apart and averaged LD values in distance classes of 1 kb for lineages 1 and 4 and 10 kb for lineages 2 and 3, to minimize noise due to low genetic diversity. Only sites without missing data and with a minor allele frequency above 10% were included, to minimize the dependence of r^2 on minor allele frequency (65). Recombination rates were estimated for each chromosome with Pairwise in LDhat version 2.2 (66). Singletons and sites with missing data were excluded.

Pathogenicity tests. We used pathotyping data for 31 isolates previously described by Gallet et al. (38). We supplemented this data set with pathotyping data for 27 isolates produced by the same authors, using the same protocol but not included in their publication due to uncertainty in the nature of the rice subspecies of origin. We used a combination of multilocus microsatellite and SNP data to assign the 58 pathotyped isolates to the six lineages, because SNP data were available for only 30 pathotyped isolates (20 of the 31 isolates from Gallet et al. and 10 of the 27 additional isolates). Multilocus microsatellite genotypes at 12 loci were obtained from the Saleh et al. (21) data set or were produced as described by Saleh et al. (21). We improved the accuracy of assignment tests by adding to the full data set the 19 isolates that had been sequenced but for which no pathotyping data were available, which included 77 multilocus genotypes in total (58 pathotyped isolates and 19 additional nonpathotyped isolates). For 49

of the 77 isolates for which genomic data were available, we retained 1% of the SNP loci with no missing data (i.e., 164 SNPs). Missing data were introduced at SNP and microsatellite loci for the 28 nonsequenced isolates and the four sequenced isolates without microsatellite data.

The Structure 2.3.1 program was used for determining assignments (67–69). The model implemented allowed admixture and correlation in allele frequencies. Burn-in length was set at 10,000 iterations, and the burn-in period was followed by 40,000 iterations. Four independent runs were performed to check for convergence. At $K = 6$, the four main clusters identified with the full genomic data set were recovered, although 15 of the 77 genotypes could not be assigned due to admixture or a lack of power. Finally, 46 of the 58 isolates inoculated could be assigned to lineages 1 to 4; the other 12 isolates could not be assigned to a specific lineage among lineages 1, 5, and 6 and were not analyzed further (Fig. S4). Infection success was analyzed with a generalized linear model with a binomial error structure and logit link function. Treatment contrasts were used to assess the specific degrees of freedom of main effects and interactions.

Genome scan for genetic exchanges. Probabilistic chromosome painting was carried out with Chromopainter version 0.0.4 (70). This method "paints" individuals in "recipient" populations as a combination of segments from "donor" populations, using linkage information for probability computation and assuming that linked alleles are more likely to be exchanged together during recombination events. All lineages were used as donors, but only lineages 1 to 4 were used as recipients (sample sizes were too small for lineages 5 and 6). We initially ran the model using increments of 50 expectation-maximization iterations, starting at 10 iterations, and we examined the convergence of parameter estimates to determine how many iterations to use. Hence, the recombination scaling constant N_e and emission probabilities (μ) were estimated in lineages 1 to 4 by running the expectation-maximization algorithm with 200 iterations for each lineage and chromosome. Estimates of N_e and μ were then calculated as averages weighted by chromosome length ($N_e = 8{,}160$ for all lineages; lineage 1, $\mu = 0.0000506$; lineage 2, $\mu = 0.0000171$; lineage 3, $\mu = 0.000021$; lineage 4, $\mu = 0.000011$). These parameter values and the per chromosome recombination rates estimated determined with LDhat were then used to paint the chromosome of each lineage, considering the remaining lineages as donors and using 200 expectation-maximization iterations. We used a probability threshold of 0.9 to assign mutations in a recipient lineage to a donor lineage.

Tip-calibrated phylogenetic analysis. Tip-calibrated phylogenetic inferences were performed with only the 48 isolates for which sampling date were recorded, i.e., all isolates except the reference strain 70-15 and strain PH0018, with the exclusion of missing data. We investigated whether the signal obtained with our data set was sufficiently high for thorough tip-dating inferences by building a phylogenetic tree with PhyML (71), without constraining tip heights on the basis of isolate sampling time, and then fitting root-to-tip distances (a proxy for the number of substitutions accumulated since the most recent common ancestor [TMRCA]) to collection dates with TempEst (70). We observed a significant positive correlation (Fig. S3), demonstrating that the temporal signal was sufficiently strong for thorough tip-dating inferences at this evolutionary scale. The tip-calibrated inferences were then carried out using Markov chain-Monte Carlo sampling in beast 1.8.2 (72). The topology was fixed as the total-evidence genome genealogy inferred with RAxML. We used an annotation of the SNPs with SNPEff (57) to partition Bayesian inference (i.e., several substitution models and rates of evolution were fitted to the different sets of SNPs during a single analysis). The optimal partitioning scheme and the best-fit nucleotide substitution model for each partitioning of the genome were estimated with PartitionFinder software (73). The best partitioning was obtained for $K = 3$ schemes (synonymous: HKY, non-synonymous: GTR and non-exonic SNPs: GTR) and was used for subsequent analyses. Node age was then estimated with this optimal partitioning scheme. Rate variation between sites was modeled with a discrete gamma distribution, with four rate categories. We assumed an uncorrelated lognormal relaxed clock, to account for rate variation between lineages. We minimized prior assumptions about demographic history, by adopting an extended Bayesian skyline plot approach, to integrate data over different coalescent histories. The tree was calibrated using tip-dates only. We applied flat priors (i.e., uniform distributions) for substitution rate ($1 \times 10^{-12} - 1 \times 10^{-2}$ substitutions/site/year) and for the age of any internal node in the tree (including the root). We ran five independent chains, in which samples were drawn every 5,000 MCMC steps, from a total of 50,000,000 steps, after a discarded burn-in of 5,000,000 steps. We checked for convergence to the stationary distribution and for sufficient sampling and mixing by inspecting posterior samples (effective sample size, >200). Parameter estimation was based on samples combined from the different chains. The best-supported tree was estimated from the combined samples and using the maximum clade credibility method implemented in TreeAnnotator.

Functional enrichment. Gene enrichment analysis was conducted with the R package TopGO for GO terms and Fisher's exact test for enrichment in HET domain genes, NLRs, small secreted protein genes, and MAX-effector genes. MAX-effector genes were those reported by de Guillen et al. (71), NLRs were those identified by Dyrka et al. (46), and small secreted proteins and HET domain proteins were identified with Ensembl's Biomart.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.01806-17.

FIG S1, DOCX file, 0.8 MB.

FIG S2, DOCX file, 0.2 MB.

FIG S3, PDF file, 0.1 MB.
FIG S4, DOCX file, 0.7 MB.
TABLE S1, XLSX file, 0.02 MB.
TABLE S2, DOCX file, 0.5 MB.
TABLE S3, XLSX file, 0.04 MB.
TABLE S4, XLSX file, 0.04 MB.
TABLE S5, XLSX file, 0.04 MB.
TABLE S6, XLSX file, 0.04 MB.

ACKNOWLEDGMENTS

We thank the scholars who contributed samples, François Bonnot and Romain Gallet for assistance with statistics, and the South Green and Migale computing facilities.

REFERENCES

1. Taylor JW, Jacobson D, Fisher M. 1999. The evolution of asexual fungi: reproduction, speciation and classification. Annu Rev Phytopathol 37:197–246. https://doi.org/10.1146/annurev.phyto.37.1.197.
2. Gladieux P, Feurtey A, Hood ME, Snirc A, Clavel J, Dutech C, Roy M, Giraud T. 2015. The population biology of fungal invasions. Mol Ecol 24:1969–1986. https://doi.org/10.1111/mec.13028.
3. Martin MD, Vieira FG, Ho SYW, Wales N, Schubert M, Seguin-Orlando A, Ristaino JB, Gilbert MT. 2016. Genomic characterization of a South American Phytophthora hybrid mandates reassessment of the geographic origins of Phytophthora infestans. Mol Biol Evol 33:478–491. https://doi.org/10.1093/molbev/msv241.
4. Simwami SP, Khayhan K, Henk DA, Aanensen DM, Boekhout T, Hagen F, Brouwer AE, Harrison TS, Donnelly CA, Fisher MC. 2011. Low diversity Cryptococcus neoformans variety grubii multilocus sequence types from Thailand are consistent with an ancestral African origin. PLoS Pathog 7:e1001343. https://doi.org/10.1371/journal.ppat.1001343.
5. Ali S, Gladieux P, Rahman H, Saqib MS, Fiaz M, Ahmad H, Leconte M, Gautier A, Justesen AF, Hovmøller MS, Enjalbert J, de Vallavieille-Pope C. 2014. Inferring the contribution of sexual reproduction, migration and off-season survival to the temporal maintenance of microbial populations: a case study on the wheat fungal pathogen Puccinia striiformis f. sp. tritici. Mol Ecol 23:603–617. https://doi.org/10.1111/mec.12629.
6. Savary S, Willocquet L, Elazegui FA, Castilla NP, Teng PS. 2000. Rice pest constraints in tropical Asia: quantification of yield losses due to rice pests in a range of production situations. Plant Dis 84:357–369. https://doi.org/10.1094/PDIS.2000.84.3.357.
7. Talbot NJ. 2003. On the trail of a cereal killer: exploring the biology of Magnaporthe grisea. Annu Rev Microbiol 57:177–202. https://doi.org/10.1146/annurev.micro.57.030502.090957.
8. Gurr S, Samalova M, Fisher M. 2011. The rise and rise of emerging infectious fungi challenges food security and ecosystem health. Fungal Biol Rev 25:181–188. https://doi.org/10.1016/j.fbr.2011.10.004.
9. Valent B. 1990. Rice blast as a model system for plant pathology. Phytopathology 80:33–36. https://doi.org/10.1094/phyto-80-33.
10. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeyer C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW. 2005. The genome sequence of the rice blast fungus Magnaporthe grisea. Nature 434:980–986. https://doi.org/10.1038/nature03449.
11. Ebbole DJ. 2007. Magnaporthe as a model for understanding host-pathogen interactions. Annu Rev Phytopathol 45:437–456. https://doi.org/10.1146/annurev.phyto.45.062806.094346.
12. Chiapello H, Mallet L, Guérin C, Aguileta G, Amselem J, Kroj T, Ortega-Abboud E, Lebrun MH, Henrissat B, Gendrault A, Rodolphe F, Tharreau D, Fournier E. 2015. Deciphering genome content and evolutionary relationships of isolates from the fungus Magnaporthe oryzae attacking different host plants. Genome Biol Evol 7:2896–2912. https://doi.org/10.1093/gbe/evv187.
13. Islam MT, Croll D, Gladieux P, Soanes DM, Persoons A, Bhattacharjee P, Hossain MS, Gupta DR, Rahman MM, Mahboob MG, Cook N, Salam MU, Surovy MZ, Sancho VB, Maciel JL, NhaniJúnior A, Castroagudín VL, Reges JT, Ceresini PC, Ravel S, Kellner R, Fournier E, Tharreau D, Lebrun MH, McDonald BA, Stitt T, Swan D, Talbot NJ, Saunders DG, Win J, Kamoun S. 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae. BMC Biol 14:84. https://doi.org/10.1186/s12915-016-0309-7.
14. Yoshida K, Saunders DGO, Mitsuoka C, Natsume S, Kosugi S, Saitoh H, Inoue Y, Chuma I, Tosa Y, Cano LM, Kamoun S, Terauchi R. 2016. Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable elements. BMC Genomics 17:370. https://doi.org/10.1186/s12864-016-2690-6.
15. Couch BC, Fudal I, Lebrun MH, Tharreau D, Valent B, van Kim P, Notteghem JL, Kohn LM. 2005. Origins of host-specific populations of the blast pathogen Magnaporthe oryzae in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics 170:613–630. https://doi.org/10.1534/genetics.105.041780.
16. Fuller DQ, Sato Y-I, Castillo C, Qin L, Weisskopf AR, Kingwell-Banham EJ, Song J, Ahn S, van Etten J. 2010. Consilience of genetics and archaeobotany in the entangled history of rice. Archaeol Anthropol Sci 2:115–131. https://doi.org/10.1007/s12520-010-0035-y.
17. Diao X, Jia G. 2017. Origin and domestication of foxtail millet, p 61–72. *In* Doust A, Diao X (ed), Genetics and genomics of Setaria. Springer, Cham, Switzerland.
18. Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, Guo Y, Lu Y, Zhou C, Fan D, Weng Q, Zhu C, Huang T, Zhang L, Wang Y, Feng L, Furuumi H, Kubo T, Miyabayashi T, Yuan X, Xu Q, Dong G, Zhan Q, Li C, Fujiyama A, Toyoda A, Lu T, Feng Q, Qian Q, Li J, Han B. 2012. A map of rice genome variation reveals the origin of cultivated rice. Nature 490:497–501. https://doi.org/10.1038/nature11532.
19. Huang X, Han B. 2015. Rice domestication occurred through single origin and multiple introgressions. Nat Plants 2:15207. https://doi.org/10.1038/nplants.2015.207.
20. Choi JY, Platts AE, Fuller DQ, Hsing YI, Wing RA, Purugganan MD. 2017. The rice paradox: multiple origins but single domestication in Asian rice. Mol Biol Evol 34:969–979. https://doi.org/10.1093/molbev/msx049.
21. Saleh D, Milazzo J, Adreit H, Fournier E, Tharreau D. 2014. South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, Magnaporthe oryzae. New Phytol 201:1440–1456. https://doi.org/10.1111/nph.12627.
22. Zeigler RS. 1998. Recombination in *Magnaporthe Grisea*. Annu Rev Phytopathol 36:249–275. https://doi.org/10.1146/annurev.phyto.36.1.249..
23. Saleh D, Xu P, Shen Y, Li C, Adreit H, Milazzo J, Ravigné V, Bazin E, Nottéghem JL, Fournier E, Tharreau D. 2012. Sex at the origin: an Asian population of the rice blast fungus Magnaporthe oryzae reproduces sexually. Mol Ecol 21:1330–1344. https://doi.org/10.1111/j.1365-294X.2012.05469.x.
24. Lu J, Tang T, Tang H, Huang J, Shi S, Wu CI. 2006. The accumulation of deleterious mutations in rice genomes: a hypothesis on the cost of domestication. Trends Genet 22:126–131. https://doi.org/10.1016/j.tig.2006.01.004.
25. Glémin S, Bataillon T. 2009. A comparative view of the evolution of grasses under domestication. New Phytol 183:273–290. https://doi.org/10.1111/j.1469-8137.2009.02884.x.
26. Stukenbrock EH, Bataillon T, Dutheil JY, Hansen TT, Li R, Zala M, McDonald BA, Wang J, Schierup MH. 2011. The making of a new pathogen:

insights from comparative population genomics of the domesticated wheat pathogen Mycosphaerella graminicola and its wild sister species. Genome Res 21:2157–2166. https://doi.org/10.1101/gr.118851.110.
27. Stukenbrock EH, Bataillon T. 2012. A population genomics perspective on the emergence and adaptation of new plant pathogens in agro-ecosystems. PLoS Pathog 8:e1002893. https://doi.org/10.1371/journal.ppat.1002893.
28. Felsenstein J. 1974. The evolutionary advantage of recombination. Genetics 78:737–756.
29. Karasov T, Messer PW, Petrov DA. 2010. Evidence that adaptation in Drosophila is not limited by mutation at single sites. PLoS Genet 6:e1000924. https://doi.org/10.1371/journal.pgen.1000924.
30. Ellison CE, Hall C, Kowbel D, Welch J, Brem RB, Glass NL, Taylor JW. 2011. Population genomics and local adaptation in wild isolates of a model microbial eukaryote. Proc Natl Acad Sci U S A 108:2831–2836. https://doi.org/10.1073/pnas.1014971108.
31. Roper M, Ellison C, Taylor JW, Glass NL. 2011. Nuclear and genome dynamics in multinucleate ascomycete fungi. Curr Biol 21:R786–R793. https://doi.org/10.1016/j.cub.2011.06.042.
32. Cheeseman K, Ropars J, Renault P, Dupont J, Gouzy J, Branca A, Abraham AL, Ceppi M, Conseiller E, Debuchy R, Malagnac F, Goarin A, Silar P, Lacoste S, Sallet E, Bensimon A, Giraud T, Brygoo Y. 2014. Multiple recent horizontal transfers of a large genomic region in cheese making fungi. Nat Commun 5:2876. https://doi.org/10.1038/ncomms3876.
33. Gladieux P, Ropars J, Badouin H, Branca A, Aguileta G, De Vienne DM, Rodríguez de la Vega RC, Branco S, Giraud T. 2014. Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. Mol Ecol 23:753–773. https://doi.org/10.1111/mec.12631.
34. Noguchi MT, Yasuda N, Fujita Y. 2006. Evidence of genetic exchange by parasexual recombination and genetic analysis of pathogenicity and mating type of parasexual recombinants in rice blast fungus, Magnaporthe oryzae. Phytopathology 96:746–750. https://doi.org/10.1094/PHYTO-96-0746.
35. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.
36. Leaché AD, Banbury BL, Felsenstein J, de Oca AN, Stamatakis A. 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst Biol 64:1032–1047. https://doi.org/10.1093/sysbio/syv053.
37. Akashi H, Osada N, Ohta T. 2012. Weak selection and protein evolution. Genetics 192:15–31. https://doi.org/10.1534/genetics.112.140178.
38. Gallet R, Fontaine C, Bonnot F, Milazzo J, Tertois C, Adreit H, Ravigné V, Fournier E, Tharreau D. 2016. Evolution of compatibility range in the rice-Magnaporthe oryzae system: an uneven distribution of R genes between rice subspecies. Phytopathology 106:348–354. https://doi.org/10.1094/PHYTO-07-15-0169-R.
39. Giraud T, Gladieux P, Gavrilets S. 2010. Linking emergence of fungal plant diseases with ecological speciation. Trends Ecol Evol 25:387–395. https://doi.org/10.1016/j.tree.2010.03.006.
40. Schulze-Lefert P, Panstruga R. 2011. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci 16:117–125. https://doi.org/10.1016/j.tplants.2011.01.001.
41. Liao J, Huang H, Meusnier I, Adreit H, Ducasse A, Bonnot F, Pan L, He X, Kroj T, Fournier E, Tharreau D, Gladieux P, Morel JB. 2016. Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. eLife 5:e19377. https://doi.org/10.7554/eLife.19377.
42. Bryant D, Moulton V. 2004. Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. Mol Biol Evol 21:255–265. https://doi.org/10.1093/molbev/msh018.
43. Bruen TC, Philippe H, Bryant D. 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics 172:2665–2681. https://doi.org/10.1534/genetics.105.048975.
44. Maddison WP, Maddison DR. 2016. Mesquite: a modular system for evolutionary analysis, version 3.11. http://mesquiteproject.org.
45. Ardlie KG, Kruglyak L, Seielstad M. 2002. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3:299–309. https://doi.org/10.1038/nrg777.
46. Dyrka W, Lamacchia M, Durrens P, Kobe B, Daskalov A, Paoletti M, Sherman DJ, Saupe SJ. 2014. Diversity and variability of NOD-like receptors in fungi. Genome Biol Evol 6:3137–3158. https://doi.org/10.1093/gbe/evu251.
47. Saupe SJ. 2000. Molecular genetics of heterokaryon incompatibility in filamentous ascomycetes. Microbiol Mol Biol Rev 64:489–502. https://doi.org/10.1128/MMBR.64.3.489-502.2000.
48. Rieux A, Balloux F. 2016. Inferences from tip-calibrated phylogenies: a review and a practical guide. Mol Ecol 25:1911–1924. https://doi.org/10.1111/mec.13586.
49. Zhao Z. 2011. New archaeobotanic data for the study of the origins of agriculture in China. Curr Anthropol 52:S295–S306. https://doi.org/10.1086/659308.
50. Castillo CC, Bellina B, Fuller DQ. 2016. Rice, beans and trade crops on the early maritime Silk Route in Southeast Asia. Antiquity 90:1255–1269. https://doi.org/10.15184/aqy.2016.175.
51. Fujiwara H. 1996. Search for the origin of rice cultivation: the ancient rice cultivation in paddy fields at the Cao Xie Shan Site in China. Society for Scientific Studies on Cultural Property, Miyazaki, Japan.
52. Fuller DQ, Qin L. 2009. Water management and labour in the origins and dispersal of Asian rice. World Archaeol 41:88–111. https://doi.org/10.1080/00438240802668321.
53. Siddiqui IH. 2008. Water works and irrigation system in India during pre-Mughal times, p 429–454. Brill Online, Leiden, The Netherlands.
54. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.
55. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
56. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. https://doi.org/10.1101/gr.107524.110.
57. Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv arXiv:12073907. [q-bio.GN]. https://arxiv.org/abs/1207.3907.
58. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w^{1118};iso-2; iso-3. Fly 6:80–92. https://doi.org/10.4161/fly.19695.
59. Notteghem JL, Silue D. 1992. Distribution of the mating type alleles in Magnaporthe grisea populations pathogenic on rice. Phytopathology 82:421–424. https://doi.org/10.1094/Phyto-82-421.
60. Jombart T, Ahmed I. 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521.
61. De Mita S, Siol M. 2012. EggLib: processing, analysis and simulation tools for population genetics and genomics. BMC Genet 13:27. https://doi.org/10.1186/1471-2156-13-27.
62. Thornton K. 2003. Libsequence: a C++ class library for evolutionary genetic analysis. Bioinformatics 19:2325–2327. https://doi.org/10.1093/bioinformatics/btg316.
63. Hill WG, Robertson A. 1968. Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231. https://doi.org/10.1007/BF01245622.
64. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
65. Amaral AJ, Megens HJ, Crooijmans RPMA, Heuven HCM, Groenen MAM. 2008. Linkage disequilibrium decay and haplotype block structure in the pig. Genetics 179:569–579. https://doi.org/10.1534/genetics.107.084277.
66. Auton A, McVean G. 2007. Recombination rate estimation in the presence of hotspots. Genome Res 17:1219–1227. https://doi.org/10.1101/gr.6386707.
67. Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.
68. Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587.
69. Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x.
70. Lawson DJ, Hellenthal G, Myers S, Falush D. 2012. Inference of popula-

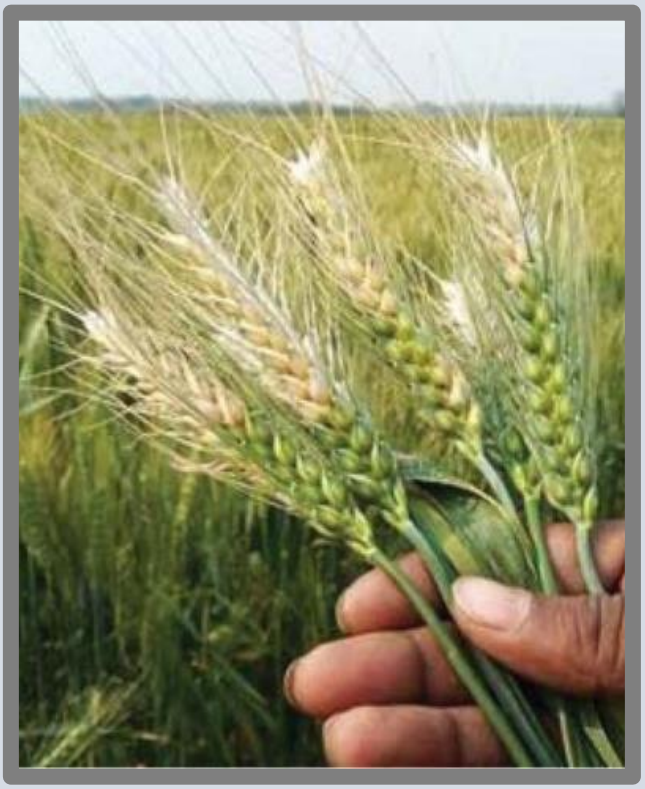
tion structure using dense haplotype data. PLoS Genet 8:e1002453. https://doi.org/10.1371/journal.pgen.1002453.

71. de Guillen K, Ortiz-Vallejo D, Gracy J, Fournier E, Kroj T, Padilla A. 2015. Structure analysis uncovers a highly diverse but structurally conserved effector family in phytopathogenic fungi. PLoS Pathog 11:e1005228. https://doi.org/10.1371/journal.ppat.1005228.
72. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973.
73. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol 34: 772–773. https://doi.org/10.1093/molbev/msw260.

CHAPITRE 1

-

Détection de la lignée Triticum causant la pyriculariose du blé



1- Contexte de l'étude

La pyriculariose du blé est une forte menace pour la sécurité alimentaire mondiale. Depuis son émergence en 1985 au Brésil dans l'état de Paraná les dégâts recensés de cette maladie sont considérables. En 1987, les pertes de rendement enregistrées dans les états brésiliens de Paraná, Mato Grosso do Sul et Sao Paulo variaient entre 10.5 et 53% (Kohli et al., 2011). Sa propagation les années suivantes dans les pays voisins n'a fait que plus de dégâts. En 1996, l'épidémie déclarée en Bolivie a entrainé près de 80% de pertes de rendement (Barea and Toledo, 1996). 70% de pertes de rendement dus à cette maladie ont été enregistrés au Paraguay en 2002 dans les champs semés tôt dans la saison (Viedma and Morel, 2002). De plus, les grains produits n'ont pas atteint les normes de qualité nécessaires pour la consommation humaine et ont été majoritairement employés pour l'alimentation animale (Kohli et al., 2011). Les premières épidémies ont été reportées en Argentine en 2007 (Alberione, 2008).

La commercialisation de semences infectées est très certainement responsable de la première apparition de la maladie en dehors du continent Sud-Américain, au Bangladesh en 2016 (Islam et al., 2016). Huit districts ont été atteints pendant cette première année d'épidémie affectant 99 259 hectares et huit autres districts supplémentaires pendant les deux années suivantes (Islam et al., 2016, 2019). Les mesures mises en place pour tenter d'éradiquer la maladie ont également fortement contribué à la diminution de la production de blé des régions impactées. En effet, le gouvernement du Bangladesh a fortement encouragé les agriculteurs des zones impactées à changer de production (maïs, sorgho, lentille …) entrainant une réduction de 52% de la production de blé dans ces districts (Islam et al., 2016). Le gouvernement Indien (pays voisin) a également banni la production de blé le long de la frontière avec le Bangladesh (~2200km) et restreint le mouvement de semences en provenance de ces régions (Islam et al., 2019).

Au vu des conséquences de la maladie, des mesures préventives pour éviter l'introduction de l'agent pathogène dans de nouveaux pays sont essentielles. Cette prévention est principalement basée sur un contrôle du matériel végétal mis en circulation et nécessite donc la mise au point de tests de détection extrêmement fiables. Un premier test de détection utilisant la technique de PCR en temps réel a été proposé en 2017 (Pieck et al., 2017), suivi d'un test de détection utilisant la technique d'amplification isothermale LAMP (Yasuhara-Bell et al., 2018). Ces deux tests ciblent la même région du génome, le locus MoT3, détectée par une approche de comparaison de génomes (Pieck et al., 2017). Toutefois le locus MoT3 est absent du génome de certains isolats responsables de la maladie entrainant des résultats faux-négatifs avec ces tests (Gupta et al., 2019).

La difficulté de la mise au point d'un test de détection moléculaire des isolats responsables de la pyriculariose du blé vient du fait que la maladie est causée par un sous-groupe d'isolats au sein de l'espèce *P. oryzae*. Ces isolats sont regroupés au sein d'une même lignée, la lignée Triticum (Gladieux et al., 2018b). Les isolats responsables de la pyriculariose du blé sont donc extrêmement proches génétiquement des autres isolats de la même espèce ne causant pas d'épidémies sur blé (<1% de divergence entre les lignées). De plus, des échanges de matériel génétique entre lignées ont été démontrés (Gladieux et al., 2018b). La différenciation des lignées hôte-spécifiques de *P. oryzae* correspond certainement à un évènement de spéciation initié mais non abouti qui se placerait dans la « zone grise » de la spéciation. Ce processus de divergence en cours a pour conséquence que les différences génétiques ne sont pas toutes fixées le long du génome. Cette structure génétique rend très difficile la recherche de polymorphismes spécifiques de la lignée Triticum qui pourraient être ciblés par des tests de détection moléculaire. La distinction entre les lignées existant au sein de *P. oryzae* est pourtant essentielle car plusieurs lignées n'infectant pas le blé ont déjà été introduites dans un grand nombre de pays non touchés par la pyriculariose du blé et peuvent, quand les conditions sont extrêmement favorables à l'infection, causer des infections opportunistes sur blé sans être à l'origine d'épidémies sur cette céréale (Farman et al., 2016). Ne pas réussir à distinguer les lignées au sein de *P. oryzae* expose donc à des résultats faussement positifs pouvant avoir de très grandes conséquences économiques (Magdama et al., 2019).

Ce chapitre de thèse présente deux articles dont l'objectif général est de proposer de nouveaux tests de détection de la pyriculariose du blé utilisant différentes techniques et améliorant le diagnostic de cette maladie dévastatrice.

Article 2: A genomic approach to develop a new qPCR test enabling the detection of the *Pyricularia oryzae* lineage causing wheat blast

Les principaux objectifs de cet article étaient de :

- Développer une approche bio-informatique basée sur la comparaison de 71 génomes de *P. oryzae* afin d'identifier des polymorphismes spécifiques de la lignée Triticum ;
- Développer un test de diagnostic par qPCR ciblant ces polymorphismes et amélioration de la capacité à détecter l'ensemble des isolats de la lignée Triticum ;
- Evaluer et valider les performances de ce test (inclusivité, spécificité, sensibilité, transférabilité et robustesse).

Plant Disease • 2019 • 0:1-11 • https://doi.org/10.1094/PDIS-04-19-0685-RE

e-Xtra*

A Genomic Approach to Develop a New qPCR Test Enabling Detection of the *Pyricularia oryzae* Lineage Causing Wheat Blast

Q:1

Maud Thierry,[1,2] Pierre Gladieux,[1] Elisabeth Fournier,[1] Didier Tharreau,[1,2] and Renaud Ioos[3,†]

[1] UMR BGPI, Montpellier University, INRA, CIRAD, Montpellier SupAgro, Montpellier, France
[2] CIRAD, UMR BGPI, F-34398 Montpellier, France
[3] ANSES Plant Health Laboratory, Mycology Unit, Domaine de Pixérécourt, Bâtiment E, F-54220 Malzéville, France

Abstract

Rapid detection is key to managing emerging diseases because it allows their spread around the world to be monitored and limited. The first major wheat blast epidemics were reported in 1985 in the Brazilian state of Paraná. Following this outbreak, the disease quickly spread to neighboring regions and countries and, in 2016, the first report of wheat blast disease outside South America was released. This Asian outbreak was due to the trade of infected South American seed, demonstrating the importance of detection tests in order to avoid importing contaminated biological material into regions free from the pathogen. Genomic analysis has revealed that one particular lineage within the fungal species *Pyricularia oryzae* is associated with this disease: the *Triticum* lineage. A comparison of 81 *Pyricularia* genomes highlighted polymorphisms specific to the *Triticum* lineage, and this study developed a real-time PCR test targeting one of these polymorphisms. The test's performance was then evaluated in order to measure its analytical specificity, analytical sensitivity, and robustness. The C17 q quantitative PCR test detected isolates belonging to the *Triticum* lineage with high sensitivity, down to 13 plasmid copies or 1 pg of genomic DNA per reaction tube. The blast-based approach developed here to study *P. oryzae* can be transposed to other emerging diseases.

Keywords: cereals and grains, emerging disease, field crops, fungi, genome comparison, infraspecific detection, pathogen detection, *Pyricularia oryzae*, wheat blast

Emerging infectious diseases (EIDs) are defined as diseases that have recently appeared on a new host or in a new area, or whose incidence has rapidly increased (Morse 2001). Various infectious disease agents and hosts may be involved in EIDs. For example, among the EIDs that have recently made the headlines, the *Zika virus* can cause fetal microcephaly in human populations (Garcez et al. 2016), the bacterium *Xylella fastidiosa* is responsible for Olive Quick Decline Syndrome (OQDS) in olive trees (Loconsole et al. 2014), and the fungus *Geomyces destructans* fulfills Koch's postulates for White-Nose Syndrome (WNS), which leads to mass bat mortalities (Fisher et al. 2012). Although these examples belong to different kingdoms and have highly contrasting biological characteristics, they have in common a sudden outbreak in a limited geographical area: the Zika outbreak occurred in 2007 in Micronesia (Ioos et al. 2014), OQDS appeared suddenly in 2010 in Italy (Loconsole et al. 2014), and WNS was described for the first time in March 2007 in New York State (Fisher et al. 2012). Therefore, avoiding the spread of EIDs as soon as possible after their outbreak is essential. To reach this goal, accurate and early diagnosis of the causal agent is a key element in EID management. It allows close monitoring of the pathogen's spread through epidemiological surveillance, thereby enabling a rapid reaction if the pathogen is introduced into novel geographical areas or preventing the trade and exchange of contaminated biological materials whenever possible (detection at points of entry and quarantine measures, for example).

Pyricularia oryzae Cavara (synonym: *Magnaporthe oryzae*) is a phytopathogenic, hemibiotrophic ascomycete fungus responsible for blast disease in several species of the family Poaceae. Blast symptoms include necrotic lesions on an infected plant's leaves, panicle necks, stems, and grains (Asuyama 1963; Te Beest et al. 2007). *P. oryzae* encompasses multiple genetic lineages that, together, affect a broad range of Poaceae hosts. Blast has been recorded on more than 25 hosts (crops or wild species) such as rice (*Oryza sativa*), wheat (*Triticum aestivum* and *Triticum durum*), foxtail millet (*Setaria italica*), or perennial ryegrass (*Lolium perenne*) (Asuyama 1963; Klaubauf et al. 2014; Maciel 2011; Milazzo et al. 2019). However, spore inoculation experiments have shown that, individually, *P. oryzae* isolates are only pathogenic on a limited range of hosts and the species contains multiple host specificity groups (Couch et al. 2005; Kato et al. 2000; Tosa et al. 2004). Population genetics and phylogenomic analyses of *P. oryzae* isolates collected from different species of Poaceae have revealed that host specificity groups defined from pathogenicity data actually represent multiple divergent lineages within the species. Each lineage is associated with a limited range of specific hosts (Gladieux et al. 2018). For instance, almost all isolates sampled on wheat belong to the same host-specific lineage: the *Triticum* lineage. Genetic divergence between host-specific lineages is very low (less than 1% genome-wide pairwise distance, measured as the number of nucleotide differences per kilobase) and genetic exchanges have recently occurred between lineages (Gladieux et al. 2018). However, host specificity is not strict, with some isolates capable of infecting host genera that are different from their host of origin, although generally with less severe symptoms than on the original host (Asuyama 1963; Farman et al. 2017). Some of these opportunistic infections have been well documented, such as that of isolate WBKY11, which was sampled on a single infected wheat head in Kentucky, the United States, despite a previous total absence of wheat blast epidemics in the state. WBKY11 has been described as genetically similar to local isolates infecting ryegrass (*Lolium* spp.) and belonging to the *Lolium* lineage, which is the closest host-specific lineage to that of *Triticum* (Farman et al. 2017; Gladieux et al. 2018). At the genetic level, host specificity has been linked in some cases to effector-triggered immunity. For example, almost all isolates infecting rice had lost the effector *AVR1-CO39* that encodes a protein recognized by rice resistance proteins (Cesari et al. 2013; Couch et al. 2005; Kang et al. 1995), and expression of the

†Corresponding author: R. Ioos ; Renaud.ioos@anses.fr

*The *e*-Xtra logo stands for "electronic extra" and indicates that two supplementary figures and four supplementary tables are published online.

The author(s) declare no conflict of interest.

Accepted for publication 2 August 2019.

PWL2 gene inhibits the infection of weeping lovegrass. Effector genes *PWT3* and *PWT4* have also been shown to prevent infection of wheat varieties carrying the corresponding resistance genes *Rwt3* and *Rwt4*, respectively (Inoue et al. 2017). Thus, a hypothesis for the appearance of wheat blast is the substantial use of *rwt3* wheat varieties (lacking the *Rwt3* resistance gene and, hence, susceptible to *P. oryzae* isolates carrying *PWT3* such as ryegrass isolates), followed by functional losses of the *PWT3* gene, allowing a massive outbreak among all *Rwt3*-carrying wheat varieties (Inoue et al. 2017).

Blast is an ancient fungal disease, first recorded on rice in 1637 in China and in 1704 in Japan (Ou 1985). The causal agent was isolated and named in 1891 but had certainly existed long before these dates, and the fungus is now found worldwide. However, before 1985, there were only sporadic reports of this disease among wheat crops (India in 1922, Pakistan in 1943, and Louisiana, the United States in 1973) (Cruz and Valent 2017; Malik and Khan 1944; Mcrae 1922; Rush and Carver 1973). The first wheat blast outbreak was reported in 1985 in the Brazilian state of Paraná, in which major symptoms were observed in wheat fields across six different municipalities. This first epidemic represented a major transition in wheat blast, evolving from causing episodic symptoms to major epidemics. Indeed, this outbreak was followed by a rapid spread to other regions in Brazil (northwestern São Paulo State and southern Mato Grosso do Sul) within a year (Anjos et al. 1996; Goulart 1992; Goulart et al. 2007; Igarashi et al. 1986; Picinini and Fernandes 1990). Soon afterward, the disease spread to three neighboring countries (1996 in Bolivia, 2002 in Paraguay, and 2007 in Argentina) (Cruz and Valent 2017). The first outbreak outside the center of origin was reported in Bangladesh in 2016 and attributed to the introduction of infected seed from South America (Islam et al. 2016; Malaker et al. 2016). Approximately 15,500 ha of wheat were infected by the wheat blast pathogen in Bangladesh in 2016 (Aman 2016; Islam et al. 2016). The *P. oryzae* isolates causing this recent EID are very aggressive. They greatly harm yield and grain quality, and up to 100% yield loss has been reported on susceptible wheat varieties under field conditions (Aman 2016; Cruz and Valent 2017; Islam et al. 2019, 2016).

It is of paramount importance to ensure early diagnosis of the *P. oryzae* isolates responsible for wheat blast epidemics (hereafter called "epidemic wheat blast isolates") in order to protect wheat crops (Ceresini et al. 2018). The economic impact of misidentifying wheat blast disease may be substantial. Inaccurate identification of the causal agent can lead to drastic measures such as bans on exportation or the destruction of vast numbers of uncontaminated seed. Conversely, false-negative results (i.e., failure to detect epidemic wheat blast isolates in contaminated material) can start an epidemic in a new area. Given these considerable economic risks, the diagnostic tests used for control and surveillance must fulfill strict conditions: they must discriminate between *P. oryzae* isolates at an infraspecific level in order to distinguish isolates causing epidemic wheat blast from isolates capable of causing only opportunistic wheat infections, despite the isolates sharing identical morphological features in pure culture.

The *Triticum* lineage identified through phylogenomic analyses includes almost all of the isolates sampled on wheat (Gladieux et al. 2018). In this study, a single isolate had not been sampled on wheat and yet belonged to the *Triticum* lineage (isolate P29). P29 was collected from a *Bromus* plant in a wheat field severely affected by blast and was strongly pathogenic on wheat during pathotyping experiments (Pieck et al. 2017). On the other hand, only three isolates sampled from wheat were assigned to the *Lolium* lineage. Among these three are the previously mentioned U.S. isolate WBKY11 and isolate PY5010, described by Pieck et al. (2017) as mildly pathogenic on wheat under laboratory conditions while being highly pathogenic on *Lolium* spp. The third isolate is PY86.1, for which no data about its pathogenicity on wheat are available. These elements strongly suggest that the isolates belonging to the *Triticum* lineage are responsible for wheat blast epidemics and their regrouping within the same lineage implies the existence of genomic regions specific to these isolates. Hereafter, the term "*Triticum* lineage isolates" is used to denote isolates that were genetically assigned to the *Triticum* lineage and capable of causing wheat blast epidemics (thus, synonymous with epidemic wheat blast isolates), and "wheat-borne isolates" is the term used to refer to those isolated from wheat regardless of their genetic assignment.

DNA-based detection methods such as PCR or real-time PCR are widely used to discriminate pathogens indistinguishable by morphology and have proved their effectiveness in detecting and identifying major plant pathogens, including fungi (Schena et al. 2004). At the species level, classical phylogenetic markers used in fungal taxonomy such as internal transcribed spacers (ITS), translation elongation factors, or β-tubulins are generally relevant targets for these species-specific PCR assays. However, in cases of taxa that have recently diverged such as *P. oryzae* lineages, it is necessary to identify genomic regions with higher variability than those used at the species level (Feau et al. 2018). Furthermore, given the very low level of divergence between *P. oryzae* lineages and the existence of recent genetic exchanges between them (Gladieux et al. 2018), it is unlikely that a single test targeting one genomic region would perfectly meet all of the requirements; for instance, in terms of specificity. A recent study showed that the presence of the MoT3 locus was a discriminant feature of most isolates belonging to the *Triticum* lineage (Pieck et al. 2017). Therefore, conventional PCR, quantitative PCR (qPCR), and loop-mediated isothermal amplification oligonucleotides have been designed to specifically amplify this genomic region (Pieck et al. 2017; Yasuhara-Bell et al. 2018). However, certain isolates assigned to the *Triticum* lineage do not carry the MoT3 sequence (for example, isolate BR0032) (Gladieux et al. 2018), suggesting that some *Triticum* lineage isolates would remain undetectable despite using some of these oligonucleotides (Gupta et al. 2019; Yasuhara-Bell et al. 2019). This pitfall indicates the need to develop a complementary test targeting a different genomic region.

In this study, we (i) searched for new genomic regions carrying DNA polymorphisms that are specific to the *Triticum* lineage isolates, focusing on alleles of genes known to be involved in pathogenicity to wheat or using a global genomic comparison of *P. oryzae* genomes; (ii) developed a real-time PCR diagnostic test that complements currently existing tools; and (iii) evaluated the specificity of this new test with *P. oryzae* isolates from a broad host range such as *P. oryzae* isolates from previously untested wild hosts such as *Eriochloa* or *Cenchrus* spp.

Materials and Methods

Fungal genome assembly. In all, 81 published *Pyricularia* genome assemblies (Gladieux et al. 2018) were compared: 4 *P. grisea* genomes, 1 *P. pennisetigena* genome, and 76 *P. oryzae* genomes. Of the latter, 20 belonged to the *Triticum* lineage, 19 had been collected from *Triticum* spp., and 1 was from *Bromus tectorum* (isolate P29) (Supplementary Table S1).

Fungal DNA extraction. DNA was extracted from a collection of 114 fungal isolates, including 110 *P. oryzae*. Of the latter, 30 had been isolated from wheat and 80 had been sampled from 21 different Poaceae genera. Some of these isolates were sequenced and genetically assigned to a host-specific lineage by Gladieux et al. (2018) (Supplementary Table S2). Upstream of this study, a population genetic analysis was conducted on 90 wheat-borne isolates collected in Brazil and Bolivia between 1989 and 2013. This analysis allowed us to maximize the genetic diversity of the wheat-borne isolates tested in this study, retaining one isolate for each multilocus genotype detected. The integrity of all 114 DNA extracts was verified by their successful amplification by PCR using ITS5-4, ITS1-4, or *RPB2* (5cR/7cR) primers (O'Donnell et al. 2007; White et al. 1990) prior to further testing.

Construction of plasmids for use as positive control. The genomic region targeted by the C17 primer set was amplified with forward and reverse primers located upstream and downstream: seqF_C17 (5′-TGC GTC CCA AGA TTT TCG TGA-3′) and seqR_C17 (5′-CTA CGC ATG GTG GGG AGA GA-3′). The amplicon was inserted in a pCR4-TOPO plasmid (Invitrogen, Carlsbad, CA, U.S.A.), which was used to transform chemically competent TOP10 cells (Invitrogen), according to the manufacturer's instructions. The

bacterial cells were then subcultured overnight at 37°C, and the plasmids were purified using a NucleoSpin Plasmid kit (Macherey-Nagel, Düren, Germany). The plasmid solution was used as a positive control for C17 real-time PCR.

Seed samples. *P. oryzae*-free wheat seed (variety Filon) from a field in Montpellier, France were used for the negative control and artificial inoculation with the pathogen. The seed were provided by C. Rudelle (GEVES, Montpellier, France). A culture of *Triticum* lineage isolate BL0044 was grown on potato dextrose agar (PDA) medium at 26°C, in a level 3 biosafety room at ANSES, Malzéville, France. A plug of mycelium was taken from the edge of the actively growing culture and used to inoculate a sample of 400 wheat grains laid over 15 ml of potato dextrose broth (PDB) in a sterile Petri dish. After 48 h of incubation, an additional 15 ml of PDB was added and the mixture was ground in a 125-ml sterilized glass bowl using a Microtron MB 550 laboratory mixer (Kinematica, Lucerne, Switzerland). A negative control seed sample was also prepared, using a sterile PDA plug for inoculation. For each seed sample, one 500-µl sample was taken from the homogenate using a truncated 1-ml pipet tip and used for total DNA extraction with NucleoSpin Plant II DNA extraction kit (Macherey-Nagel), following the manufacturer's recommendations. *Triticum* lineage was detected by real-time PCR under optimized conditions, using these raw DNA extracts as templates.

Plant varieties for pathogenicity tests. *P. oryzae* isolates were tested for pathogenicity on wheat variety Thésée, which is susceptible to wheat blast, and on a commercial mixture of three ryegrass varieties (Pronto, Fandango, and Funk) used as susceptibility checks. Compatibility between *P. oryzae* isolates and wheat or ryegrass was evaluated under controlled conditions, by inoculating wheat leaves (5 plants/isolate) or ryegrass leaves (more than 10 plants/isolate) with a conidial suspension using two protocols. First, as described for rice by Gallet et al. (2016), 10 ml of conidial suspension (20,000 conidia/ml) with 1% gelatin was sprayed on 2-week-old seedlings for each isolate. This method of inoculation is softer and is closer to natural inoculation but was not adapted to inoculate epidemic wheat blast isolates as positive controls in our biosafety facilities. Second, a suspension of 50,000 conidia/ml was applied using a brush in a delimited wheat leaf region. This method may be deemed as more invasive because brush movements can injure leaf tissues but it enabled the introduction of positive controls (isolates BR0032 and BL0028) in the compatibility experiment. In both cases, the inoculated plants were kept overnight at 27°C and 100% humidity and then for a further 7 days with alternation of day and night (13 h at 27°C and 11 h at 21°C, respectively) before noting symptoms.

Search for *PWT3* and *PWT4* alleles in *P. oryzae* genomes. To determine whether the sequences of the avirulence genes *PWT3* and *PWT4* could be used as genomic hallmarks of the *Triticum* lineage, BLASTN (version 2.6.0) was used to search for alleles of these genes in the 81 *P. oryzae* genomes (options -task blastn and e-value 0.0001). The query sequences used were the nucleotide sequences of the avirulent alleles; that is, A0 for *PWT3* (accession number LC202650.1) and Br58 for *PWT4* (accession number LC202655.1). Some of the DNA sequences located upstream or downstream of these genes were repeated multiple times in the assembled genomes, such as positions 1 to 159 and 531 to 603 for the PWT3 A0 allele, and positions 1 to 211 and 853 to 927 for the PWT4 Br58 allele. Therefore, every hit that only mapped these repeated regions was removed. In addition, when two parts of the gene were identified on the same scaffold (indicating a large insertion in the gene sequence) or in different scaffolds on the assembly (due to either a genomic rearrangement or an assembly issue), the sequences were merged. All sequences were aligned with Muscle (version 3.8.31) and compared (data not shown).

Comparison of *P. oryzae* genomes to identify DNA regions unique to the *Triticum* lineage. To detect DNA regions that were specific to the *Triticum* lineage, we designed a bioinformatics approach using PYTHON scripts and based on comparing 81 previously published assembled genomes. We focused on polymorphisms that were fixed in the *Triticum* lineage. The genome of isolate BR0032 sampled from wheat and belonging to the *Triticum* lineage was chosen as a reference. This genome was fragmented into sequences of 500 nucleotides with 250 bp overlap. To select fragments of interest, we conducted a homology analysis with other genomes using BLASTN (version 2.6.0), specifying -task blastn. Fragments were then selected if two criteria were fulfilled: (i) a sequence strictly identical to the candidate fragment was found in each of the genomes belonging to the *Triticum* lineage and (ii) no sequence 100% identical to the candidate fragment was found in any of the genomes belonging to other host-specific lineages. A first screening of the fragments was performed using only three genomes (from *Triticum*, *Lolium*, and *Oryza* lineages), and every fragment retained was then screened using the other 78 genomes. Finally, for each selected fragment, the homologous sequence in each assembled genome was extracted and all the sequences were aligned using Muscle (version 3.8.31). This study's bioinformatics analysis is summarized and sketched in Figure 1.

Development of a qPCR diagnostic test. *Primer design.* Candidate primers and probes were designed to specifically amplify DNA from the *Triticum* lineage genotypes, targeting the *Triticum* lineage-specific positions identified by bioinformatics analyses. Forward and reverse primers were designed so as to amplify a DNA fragment of 90 to 120 bp using PCR. The polymorphisms specific to the *Triticum* lineage were preferentially positioned at the 3′ end of the forward and reverse primers, in order to enhance PCR specificity. Primers and 5′ 6-*FAM*/3′ *BHQ-1* hydrolysis probes were custom synthesized by Eurogentec (Seraing, Belgium).

Screening candidate primers for inclusivity and specificity. Inclusivity (i.e., ability to target all epidemic wheat blast isolates) and specificity (i.e., ability to avoid cross-reacting with nontarget DNA) were the main criteria for primer selection. Primers were first screened for specificity and inclusivity using conventional PCR. These two parameters were assessed using DNA from a small panel consisting of seven isolates sampled on different hosts (three on ryegrass: AG0064, PL2-1, and 365; one on finger millet: IN0113; one on rice: BR0019; and two wheat-borne isolates: AG0103 and BL0093). PCRs to enable screening used the qPCR core kit No ROX (Eurogentec) with the following reaction mixture: 2 µl of DNA template (0.5 ng/µl), 1× polymerase buffer, 5 mM $MgCl_2$, 0.3 µM forward and reverse primers, 4× 0.2 mM dNTPs, HotGoldStar DNA polymerase at 0.025 U/µl, and molecular-grade water to 20 µl. The PCR conditions were 10 min at 95°C, then 40 cycles of 15 s at 95°C and 55 s at 60°C. Amplification products were resolved by ethidium bromide staining after 50 min of electrophoresis on 2% agarose gel. Q:2

Specificity and inclusivity were then further assessed by combining the best candidate-specific primer pairs with a hydrolysis probe in real-time PCR. The hydrolysis probes were designed to target polymorphisms specific to the *Triticum* lineage, whenever possible. The specificity of primer and probe combinations was tested using DNA from a larger panel of 114 isolates, including 110 *P. oryzae* isolates sampled on 21 different host genera. Real-time PCRs were carried out in a 20-µl reaction volume containing 2 µl of DNA template (0.5 ng/µl), 1× qPCR core kit No ROX (Eurogentec) polymerase buffer, 5 mM $MgCl_2$, 0.3 µM forward and reverse primers, 0.1 µM probe, 4× 0.2 mM dNTPs, HotGoldStar DNA polymerase at 0.025 U/µl, and molecular-grade water to 20 µl. The following PCR profile was used: 10 min at 95°C, then 40 cycles of 15 s at 95°C and 55 s at 60°C. The specificity and inclusivity of the MoT3 test were assessed on the same 114-isolate panel, using the PCR conditions described above.

C17-F/-R/-P specific test performance values. For the C17 primer/probe combination, the master mix composition and cycling conditions were adjusted to optimize specificity and sensitivity (mean cycle threshold [Ct] values). Optimal real-time PCR conditions were as follows: 2 µl of DNA template, 1× polymerase buffer (qPCR core kit No ROX; Eurogentec), 5 mM $MgCl_2$, 0.3 µM forward C17-F (5′-CGA TAG AAA CTT GAG GAA GAT CAA GTA AG-3′) and reverse C17-R (5′-TCA CCG AGA GAT GTG CCA C-3′) primers, 0.05 µM C17-P probe (5′-FAM- TCG CTA ACA ATG TCC ACC CCG CC-BHQ1-3′), 4× 0.2 mM dNTPs, HotGoldStar DNA polymerase at 0.025 U/µl, and molecular-grade water to 20 µl. PCR cycles were identical to those used for primer screening, except that the

hybridization or polymerization temperature was increased to 65°C. The following performance criteria were assessed: analytical specificity, analytical sensitivity, transferability, and robustness. Analytical specificity, analytical sensitivity, and robustness measures were determined with data generated by a Stratagene Mx3005P real-time thermocycler (Agilent Technologies, La Jolla, CA, U.S.A.) and extracted using Mx3005P software (V4.10) with a threshold set at 400. Analytical specificity was assessed using duplicated reactions using genomic DNA (gDNA) from the 114 *P. oryzae* isolates previously adjusted to 0.5 ng/µl. Analytical sensitivity was assessed using a 10-fold dilution series of the plasmid DNA-positive controls containing the C17-F/-R amplicon diluted in ultrapure water and with a dilution of gDNA from isolates BR0032 and BL0028. The dilution series ranged from 13×10^3 to 1.3 plasmid copies/µl and gDNA at 0.5 ng/µl to 0.05 pg/µl. The limit of detection (LOD) was determined as the minimal quantity of target DNA that could be consistently amplified by qPCR (at least 10 replicates). A standard curve was constructed based on these data.

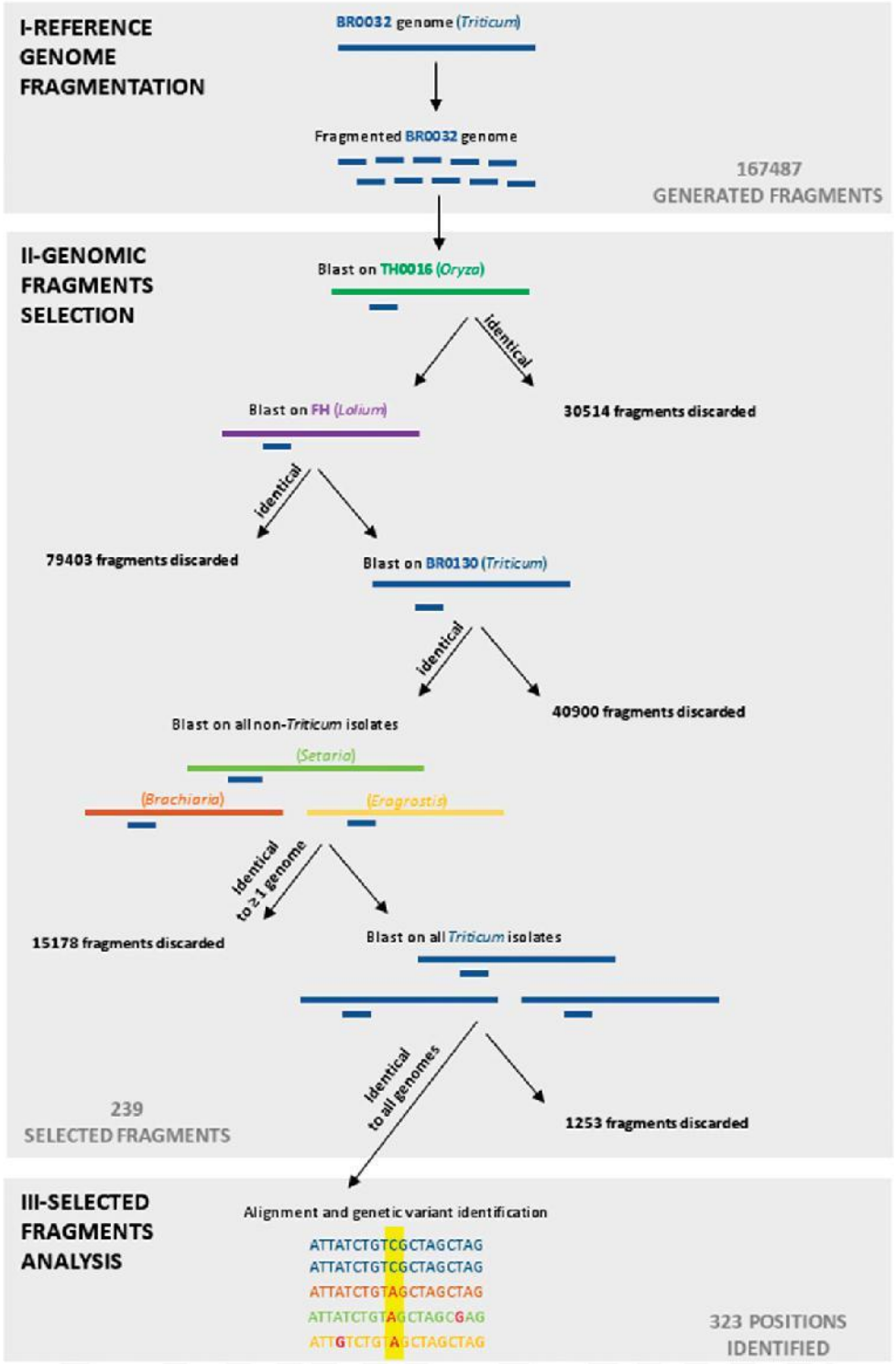


Fig. 1. Bioinformatics approach used to identify polymorphisms specific to the *Triticum* lineage of *Pyricularia oryzae*. Colors of the schematic genomes represent isolate lineages.

The test's transferability was assessed using three replicates with different brands of qPCR reagents: 2× Takyon for Probe Assay, No ROX blue (Eurogentec), LightCycler 480 Probes Master (Roche, Meylan, France), and Takyon Core Kit for Probe Assay, No ROX blue (Eurogentec), another type of thermocycler (Rotor-Gene 6500; Corbett Research, Mortlake, Australia), and a different laboratory (ANSES Plant Health Laboratory). To assess the robustness of the qPCR assay, several reaction parameters were deliberately modified and the assay was carried out using 10 replicates of different DNA templates. The qPCR assay's robustness was challenged by varying the reaction volume or DNA template volume by ±10%, and slightly varying the hybridization temperature (±2°C) to verify whether the test could withstand experimental variations without compromising sensitivity and specificity. Transferability and robustness were determined using a panel of different DNA types as templates. The panel consisted of (i) target DNA close to the LOD (10× LOD and 100× LOD plasmid DNA) and *Triticum* lineage BL0044 gDNA at 0.5 ng/µl, (ii) nontarget DNA (AG0064 ryegrass isolate at 0.5 ng/µl and IN0113 finger millet isolate gDNA), and (iii) total DNA extracted from healthy wheat seed and wheat seed artificially inoculated with BL0044.

Results

***PWT3* and *PWT4* avirulence genes are not specific markers to the *Triticum* lineage.** The *PWT3* and *PWT4* alleles (including presence or absence polymorphisms) were examined in 81 published and assembled genomes, including 76 *P. oryzae* genomes, in order to determine whether alleles of these effectors can be diagnostic of the *Triticum* lineage. The *PWT4* avirulence gene sequence was absent in 50 of the 81 genomes studied, including 14 isolates from the *Triticum* lineage and 36 isolates from other lineages. The complete and functional allele (i.e., the complete coding sequence and its upstream region) was only identified in the genome of isolate Br58, as already described by Inoue et al. (2017). Other genomes carried truncated alleles lacking various genic regions; one of these alleles was shared among isolates of *Lolium* lineage and certain isolates of the *Triticum* lineage (T25 and Br48). *PWT3* was found in all genomes except *P. oryzae* isolate EI9411 sampled on *Eleusine* spp. and *P. pennisetigena* isolate Pm1 sampled from *Pennisetum* spp. Various *PWT3* gene alleles were retrieved from these assemblies, including alleles displaying large insertions or gene fragmentation (i.e., fragments of the gene on different assembly scaffolds).

Sequence comparison, based on the regions conserved between different *PWT3* alleles, highlighted that *Triticum* lineage isolates shared alleles with non*Triticum* lineage isolates such as *Lolium* or *Stenotaphrum* lineage isolates. For both *PWT3* and *PWT4*, some alleles were shared between isolates of the *Triticum* lineage and other lineages. Therefore, *PWT3* and *PWT4* avirulence genes were not suitable for designing markers specific to the *Triticum* lineage and were not further used in developing a diagnostic test.

Comparative genomics identifies candidate genetic markers unique to the *Triticum* lineage. The genome of isolate BR0032, used as a reference for the *Triticum* lineage, was fragmented in silico into 167,487 fragments. Our bioinformatics approach retained only fragments for which (i) a strictly identical (100% identity over 500 bp) sequence was found in all genomes from the *Triticum* lineage and (ii) hits in genomes from other lineages had <100% identity (Fig. 1). In total, 239 fragments were selected and their corresponding alignments were generated. At the end of this process, 323 polymorphisms specific to the *Triticum* lineage were identified. These positions were spread over 135 fragments belonging to four distinct scaffolds of the reference genome BR0032 (Supplementary Table S3).

Design of *Triticum* lineage-specific oligonucleotide candidates. In total, 34 primer pairs (forward and reverse primers) were designed to specifically amplify the DNA of *Triticum* lineage isolates using PCR. These oligonucleotide combinations target the different regions on the BR0032 genome that carry some of the *Triticum* lineage-specific positions identified by our bioinformatics scripts. As much as possible, the regions selected (i) were physically distant from each other within the BR0032 genome and (ii) carried several polymorphisms in order to improve the oligonucleotides' specificity. Oligonucleotide sequences are given in Supplementary Table S4.

Prescreening of inclusive and specific primer sets. Among the 34 primer pairs tested, 14 (41%) yielded a stronger amplification signal with DNA from the two wheat-borne isolates included in the small panel, confirming by PCR the in silico results provided by the bioinformatics approach. However, of these 14 primer pairs, only 5 (C5, C7, C9, C17, and C19) generated no cross-reaction with DNA from the nontarget isolates in our small panel, and 4 of these (C5, C9, C17, and C19) were selected for further analyses. An associated hydrolysis probe was designed for use in combination with each selected primer pair in real-time PCR.

Final selection and evaluation of *Triticum* lineage-specific oligonucleotides. The four selected oligonucleotide combinations target sequences located in different regions of genome BR0032. Combinations C5, C9, and C17 target scaffold 15 of genome BR0032 (with an approximately 29,500-bp distance between C9 and C5 and an approximately 109,000-bp distance between C5 and C17), whereas combination C19 targets scaffold 17. The previously published primer/probe combination MoT3_1F, MotT3_1R, and MoT3_FAM2, which targets epidemic wheat blast isolates of *P. oryzae* (Pieck et al. 2017), was also included for comparison. The specificity and inclusivity of the combinations were assessed using real-time PCR, with a larger panel of 114 DNA extracts. The results showed that none of them were able to achieve full inclusivity and full specificity. In our PCR conditions, combinations MoT3 and C19 failed to yield positive results with DNA from 2 and 3 different wheat-borne isolates, respectively, out of the 30 in our panel, whereas they yielded cross-reactions with a single but different DNA extract from nontarget isolates (Fig. 2). When PCR used the three oligonucleotide combinations designed on scaffold 15 (C5, C9, and C17), it successfully amplified DNA from 100% of the 30 wheat-borne isolates tested. However, using these combinations, cross-amplifications were observed with DNA from isolates sampled on other host plants (Fig. 2). Based on these results, combination C17 was selected because, under our PCR conditions, it successfully amplified DNA from all 30 wheat-borne isolates in our panel yet yielded the lowest rate of cross-reactions (only four nontarget isolates: CR0023, CR0057, JP0033, and IN0113). After adjusting the primer mix composition and increasing the hybridization temperature to 65°C, all 30 DNA extracts from wheat-borne isolates were amplified (with Ct values ranging from 25.74 to 30.69, including isolate P29, sampled from *Bromus* spp. but belonging to the *Triticum* lineage), whereas late amplification was not observed with DNA from isolate IN0113 sampled from *Eleusine* spp. (Fig. 2).

C17-F/-R/-P test performance. *Test sensitivity.* The analytical sensitivity of test C17 was determined by analyzing three replicates of a serial 10-fold dilution in ultrapure water of gDNA of the isolates BR0032 and BL0028 or a dilution of the PCR target inserted into plasmid DNA. The test successfully amplified 100% of the replicates down to 1 pg of gDNA for both isolates and down to 13 plasmid copies/reaction tube (Fig. 3). This limit of detection was confirmed by a further test using 10 replicates. The regression coefficient (R^2) was 0.981, showing a good correlation between Ct values and the initial quantity of target DNA loaded (Fig. 3). As a proof of concept, DNA extracted from noninoculated seed and DNA extracted from seed artificially inoculated with isolate BL0044 were tested by real-time PCR using the C17-F/-R/-P set. The test consistently detected BL0044 DNA in the inoculated seed, whereas no signal was observed with the control seed sample.

C17 test robustness and transferability. When minor changes in volume and temperature were applied to the original PCR conditions, the mean Ct values obtained with the DNA templates tested were significantly modified (P value = 2.2e-16). However, Ct values measured under these different conditions are of the same order of magnitude without altering the qualitative results (+/−) and, thus, corroborate the specificity and sensitivity of the C17 test. An exception to this was when late amplification of the DNA from isolate IN0113 (Ct = 38.83) was observed for 1 of the 10 replicates when

the hybridization temperature was decreased to 63°C. The test's transferability in terms of other lab conditions was evaluated by changing the type of thermocycler, the real-time master mix brand, and the operator. In most cases, specificity and sensitivity of the C17 test were not qualitatively affected by these modifications. However, it is worth mentioning that changing the real-time PCR master mix type led to late amplifications of DNA from the IN0113 *Eleusine* isolate. Quantitatively, Ct values were generally improved (i.e., determined earlier) using the Rotorgene thermocycler. Data are presented in Figure 4.

Characterizing cross-amplified taxa. Because DNA from isolates CR0023, CR0057, and JP0033 were consistently cross-amplified by the C17-F/-R/-P test, PCR primers were designed to flank the region amplified by primer set C17 in order to

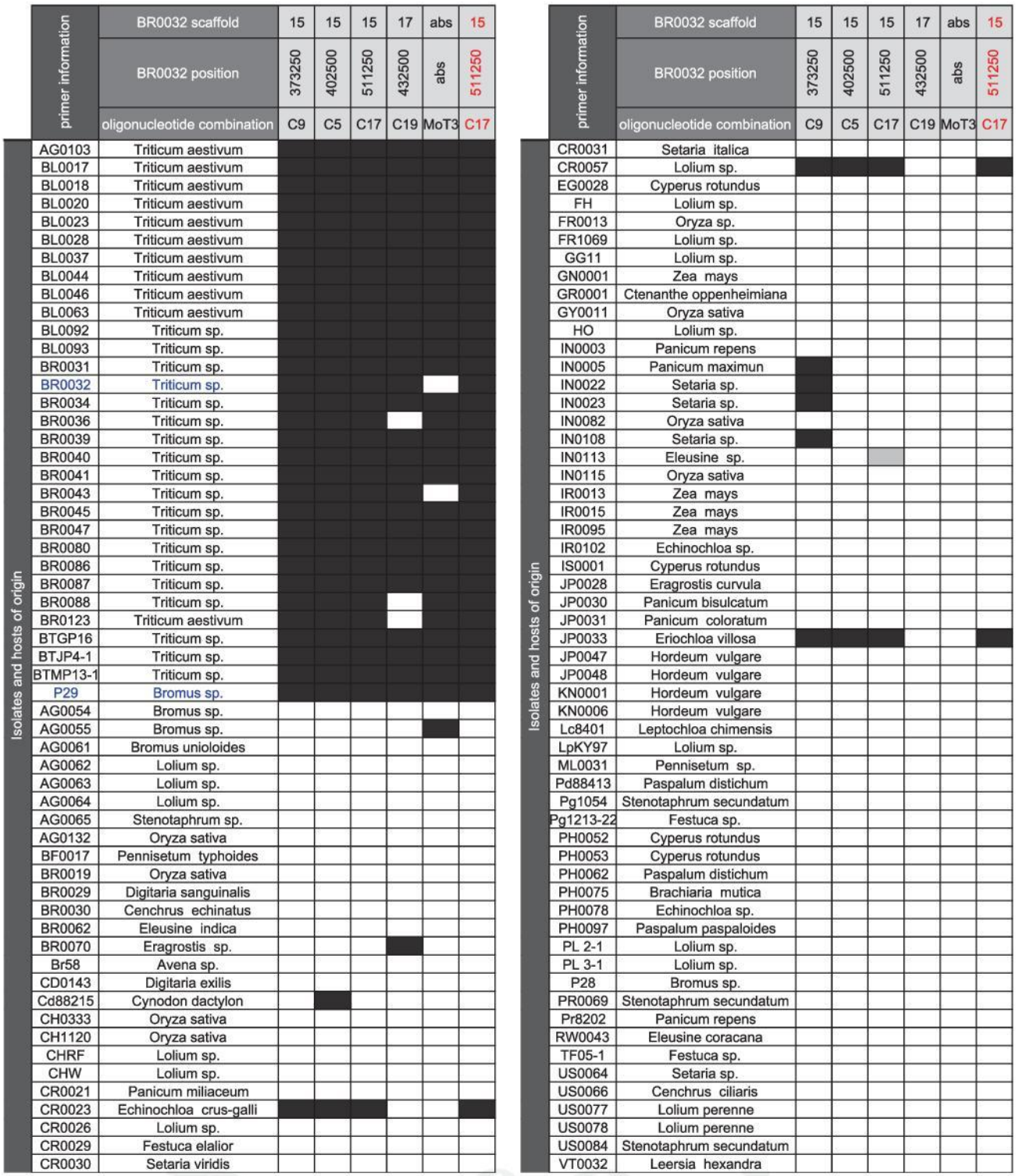


Fig. 2. List of *Pyricularia oryzae* isolates used to assess the specificity and inclusivity of the selected primer/probe combinations in real-time PCR. The color of the boxes indicates the amplification status of the 114 DNA extracts tested using primer sets C5, C9, C17, C19, and MoT3 (dark gray = amplification, light gray = late amplification, and white = no amplification). Last column (red) corresponds to the C17 primer set test using optimized quantitative PCR parameters (including higher hybridization temperature). Isolates in blue characters were genetically assigned to the *Triticum* lineage by Gladieux et al. (2018).

Sanger-sequence the entire locus. Sequencing confirmed that isolates CR0023, CR0057, and JP0033 shared 100% identity with *Triticum* lineage isolates at this locus. Within the C17-F/-R amplicon, a single-nucleotide polymorphism (SNP) was identified for isolate CR0023. Sequencing isolate IN0113 showed that the C17F primer's target region differed by two substitutions, located near the 3′ end (Supplementary Fig. S1). These substitutions lead to the appearance of a sequence that is identical to the last five bases at the 3′ end of C17's forward primer and could explain the late amplification of this DNA at a lower hybridization temperature. The pathogenicity of isolates CR0023, CR0057, and JP0033 was assessed by two methods. A brush inoculation confirmed the high leaf susceptibility of the wheat variety Thésée to wheat-borne isolates BR0032 and BL0028. Concerning non-wheat-borne isolates, except for CR0023, no reaction was observed. With isolate CR0023, limited necrosis could be observed at the margin or in the central vein of the leaves with this brush inoculation protocol for 5 of the 12 leaves inoculated. Spraying conidia on wheat cultivar leaves as well as on susceptible ryegrass varieties is a less invasive method that mimics the natural inoculation process. Using this inoculation protocol,

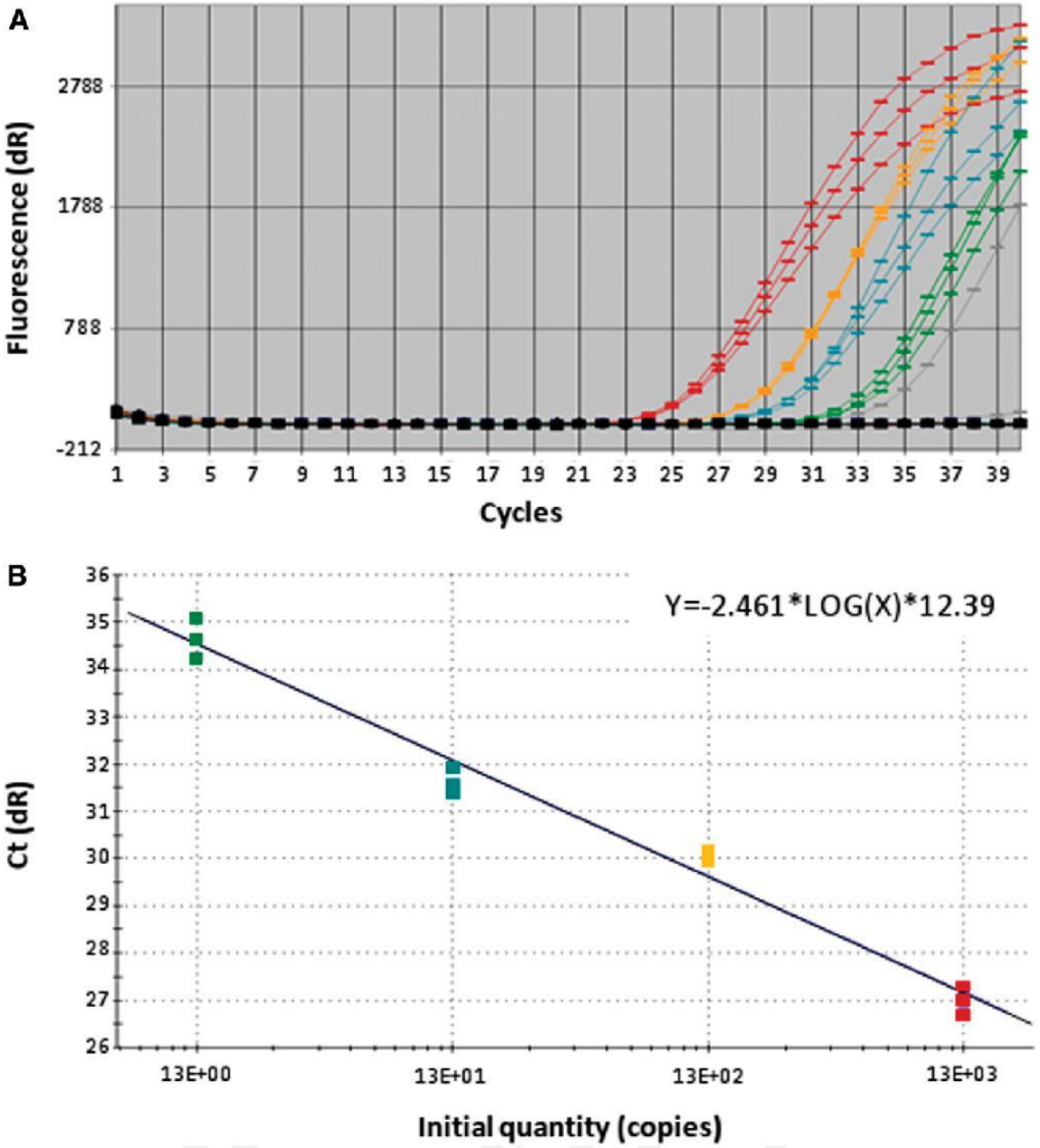


Fig. 3. Sensitivity of the primer set C17 analyzed using a 10-fold serial dilution of plasmid DNA for positive control: gray = 1.3 copies/µl, green = 13 copies/µl, blue = 13E+01 copies/µl, yellow = 13E+02 copies/µl, and red = 13E+03 copies/µl). **A,** Amplification plot of a serial plasmidic 10-fold dilution and **B,** associated standard curve. Ct = cycle threshold.

none of the non-wheat-borne isolates tested induced any symptoms on wheat leaves but they created at least one lesion on susceptible checks of ryegrass, confirming their pathogenic potential in these conditions. On ryegrass, the strongest symptoms were observed for CR0023 (sampled from *Echinochloa crus-galli*), intermediate symptoms for CR0057 (*Lolium* sp.) and only one lesion was observed for JP0033 (*Eriochloa villosa*) on ryegrass. Finally, because DNA from AG0055 isolated from *Bromus* spp. cross-reacted with the MoT3-based PCR test (Pieck et al. 2017) in our conditions, and because no genomic data are available concerning this isolate to validate the lineage it belongs to, this isolate was inoculated during the same experiment on wheat leaves (leaf brush) to validate its incompatibility with this host. This isolate was not able to induce any symptoms and, therefore, was deemed unable to infect wheat (Supplementary Fig. S2).

Discussion

Infecting a novel host is a major route of transmission in the emergence of infectious disease (Giraud et al. 2010; Gladieux et al. 2015). The management of new emerging diseases requires tools that can discriminate between emerging and established populations, in order to detect the pathogen at the earliest point and to monitor and limit its spread. Detection based on DNA amplification is a rapid and sensitive alternative to morphology-based identification and to pathogenicity testing and, today, is favored for infraspecific detection (Feau et al. 2018; Firth and Lipkin 2013). Indeed, morphology-based identification is generally not diagnostic at the infraspecific scale in fungi, whereas pathogenicity tests are resource consuming and their outcome may be unclear (for instance, if the isolates surveyed represent opportunistic infections and do not constitute a self-sustainable population on the new host). The DNA-detection method can target either known host specificity determinants or other markers specific to the subgroup of isolates infecting a particular host (Bühlmann et al. 2013; Fernandez and Orth 2018; Franco Ortega et al. 2018; Singh and Kapoor 2018; Van Dam et al. 2018). In this study, an overall genome comparison of 81 assembled genomes enabled the identification of 323 polymorphisms specific to the *Triticum* lineage of *P. oryzae,* the causal agent of wheat blast. Some of these polymorphisms were successfully exploited to develop a qPCR-based detection test that specifically targets these positions, facilitating diagnosis of the causal agent in any particular case of wheat blast. We conducted an extensive validation process to clearly identify this detection test's limits, analyzing its inclusivity, specificity, sensitivity, robustness, and transferability.

Before starting an extensive search for new genetic targets to design a diagnostic test, we set out to assess the usability of previously described determinants of host specificity as markers specific to the *Triticum* lineage of *P. oryzae*. Inoue et al. (2017) showed that functional losses of the *PWT3* and *PWT4* effectors, which trigger an immune response in wheat varieties carrying cognate resistance genes, were key determinants in the spread of *P. oryzae* as a wheat pathogen, suggesting that *PWT3* and *PWT4* alleles would be ideal genetic markers for discriminating the *Triticum* lineage from other lineages. Our study clearly showed high sequence heterogeneity of *PWT3* within the *Triticum* lineage and allele sharing between the *Triticum* lineage and other lineages. This pattern can be explained by several evolutionary scenarios, including the emergence of the *Triticum* lineage through admixture with several other lineages, incomplete lineage sorting at these genes, or gene flow into the *Triticum* lineage subsequent to its emergence. Regardless of the cause of allelic polymorphism at *PWT3* and *PWT4*, no allele was specific to the *Triticum* lineage, thus precluding the use of these genes as targets for specific markers.

As an alternative means to identify further regions specific to the *Triticum* lineage, we developed a bioinformatics approach based on comparing the genomes from *P. oryzae* isolates of different lineages. We also included four genomes from *P. grisea* isolates and one genome from *P. pennisetigena* isolates. Our approach was not designed to find all polymorphisms specific to the *Triticum* lineage but to select only those located in fixed regions of the *Triticum* lineage. In fact, the only SNPs selected were *Triticum*-specific and identified in fixed 500-bp sequences of *Triticum* lineage genomes. Interestingly, the 323 polymorphisms identified were not evenly scattered all over the genome but clustered in a limited number of regions, with 88% of polymorphisms found between positions 348,000 and 720,750 of scaffold 15 of the reference *Triticum*-lineage genome BR0032, and 11% of polymorphisms found between positions 216,750 and 595,750 of scaffold 17. In all, almost all of the positions found (99%) were concentrated in two regions accounting for less than 2% of the reference genome. These two regions may have experienced positive selection in the *Triticum* lineage, or may be linked to regions that have experienced positive selection. Therefore, the

			ROBUSTNESS						TRANSFERABILITY			
		HYBRIDIZATION T°	63	67	65	65	65	65	65	65	65	65
		DNA VOLUME	2	2	1.8	2.2	2	2	2	2	2	2
		TOTAL VOLUME	20	20	20	20	18	22	20	20	20	20
		THERMOCYCLER	A	A	A	A	A	A	B	B	B	B
		KIT	1	1	1	1	1	1	1	2	3	4
DNA	plasmid	100x T+LOD	**29.6** +-0.1	**30** +-0.3	**29.4** +-0.1	**29.5** +-0.1	**29.1** +-0.2	**29.6** +-0.2	**22.2** +-0.1	**22.7** +-0.2	**22.5** +-0.2	**22.9** +-0.1
		10x T+LOD	**31.4** +-0.3	**32.36** +-0.5	**31.9** +-0.3	**32.2** +-0.2	**30.9** +-0.2	**31.4** +-0.3	**25.8** +-0.4	**26.3** +-0.1	**25.4** +-0	**26.4** +-0.1
	genomic	(wheat) BL0044	**27.1** +-0.2	**27.3** +-0.2	**27.1** +-0.1	**26.8** +-0.2	**27** +-0.2	**26.9** +-0.1	**24.4** +-0.3	**25.8** +-0.2	**25.7** +-0.5	**24.9** +-0.3
		(eleusine) IN0113	**38.83** (1/10)	-	-	-	-	-	-	-	-	-
		(lolium) AG0064	-	-	-	-	-	-	-	-	-	-
	seeds	inoculated seeds	**31.8** +-0.2	**32.2** +-0.3	**32** +-0.1	**31.4** +-0.2	**31.6** +-0.2	**31.7** +-0.2	**25.2** +-0.2	**25.6** +-0.2	**25.8** +-0.1	**25.6** +-0.4
		non-inoculated seeds	-	-	-	-	-	-	-	-	-	-

Fig. 4. Assessment of C17 test performance measuring amplification signals under different PCR conditions by varying hybridization temperature, DNA volume, total volume of PCR mix, thermocycler (A = Stratagene Mx3005P and B = Corbett rotor-gene 6005), or PCR kit brand (1 = Core kit NO ROX, Eurogentec; 2 = 2× Takyon for Probe Assay, No ROX blue; 3 = LightCycler 480 Probes Master; and 4 = Takyon Core Kit for Probe Assay, No ROX blue).

regions under positive selection might be functionally important in terms of pathogenicity to wheat. A similar strategy was used to detect the MoT3 marker (Pieck et al. 2017) but, though both studies have investigated the same pathogen species, this study used different genomic regions. In fact, several key parameters differed between the two studies, such as the reference genome (isolate B2 versus isolate BR0032), number of genomes studied, and threshold value used for selecting genome fragments. This example highlights that these parameters strongly affect the output when performing genome comparisons.

Based on the set of candidate-specific positions identified, 34 primer pairs were designed. These primers' specificity (ability to detect only isolates of *Triticum* lineage) and inclusivity (ability to detect any isolate of *Triticum* lineage regardless of origin) was checked using DNA samples that represented a set of isolates broader than that which was used in the bioinformatics detection process. The surveyed isolates' host of origin was known but, for most of them, genetic data were not available to confirm their lineage. However, all of the isolates in our study that were sampled on wheat (referred to here as wheat-borne isolates) were obtained from wheat blast epidemics. Of the 34 initially designed primer pairs, 14 showed stronger affinity to DNA from wheat-borne isolates. This high ratio confirmed the efficiency of our bioinformatics approach in identifying positions specific to a particular lineage within a species. However, in several cases, there were very few differences between targeted sequences of wheat-borne isolates and non-wheat-borne isolates, sometimes down to one SNP. The high sequence identity likely enabled a sufficient annealing of the primers to non-wheat-borne isolate DNA. This combination would have resulted in the weak nonspecific amplifications sometimes observed. Finally, five primer pairs successfully amplified only the two DNA extracts from wheat-borne isolates in the reduced panel, and four of them were considered for the development of a diagnostic test. *Triticum* lineage-specific hydrolysis probes were designed for each of these four primer pairs. The inclusivity and specificity of these primer/probe combinations were challenged by real-time PCR reactions with DNA from a wide range of 111 *P. oryzae*, including 30 wheat-borne isolates and 81 *P. oryzae* isolates sampled on other members of the Poaceae family. We also included the previously published MoT3 test (Pieck et al. 2017) for comparison. All primer/probe combinations tested (C5, C9, C17, C19, and MoT3) proved to be highly but not 100% specific. Because nondetection of *Triticum* lineage isolates could lead to the introduction of the pathogen in disease-free regions with potentially disastrous consequences, we ultimately selected the C17-F/-R/-P set that was able to amplify DNA from 100% of the wheat-borne isolates of our panel. Surprisingly, false-positive results were observed when C17 real-time PCR was carried out with DNA from four isolates that had not been sampled on wheat. DNA from isolates CR0023, CR0057, and JP0033 was amplified with a Ct value equivalent to *Triticum* lineage isolates while that of IN0113 led to a late amplification signal. These isolates had originally been sampled on wild Poaceae spp. in Asia: CR0023 had been sampled on *Echinochloa crus-galli* in Korea, CR0057 on *Lolium* spp. in Korea, JP0033 on *Eriochloa villosa* in Japan, and IN0113 on *Eleusine* spp. in India. Sequencing the genomic region targeted by the C17-F/-R primers confirmed that CR0023, CR0057, and JP0033 carried a sequence 100% identical to that targeted for *Triticum* lineage isolate. Such patterns (i.e., isolates carrying genomic regions specific to another lineage) have previously been observed in *P. oryzae* (Gladieux et al. 2018; Inoue et al. 2017) and explained by gene flows between lineages. In our case, wheat blast epidemics have never occurred in Japan or Korea, which rules out the possibility that these isolates belong to the *Triticum* lineage or are due to recent genetic exchanges between these isolates and *Triticum* lineage isolates. Furthermore, these isolates were also positive in the C5 and C9 tests designed on the same scaffold of the BR0032 genome, suggesting that they actually share a significant genomic stretch with *Triticum* lineage isolates and making the hypothesis of a convergent evolution unlikely. Given these facts, the hypothesis of incomplete lineage sorting is favored to explain these "false positives". Pathotyping tests confirmed that these three isolates were not able to infect wheat by spraying a suspension of 20,000 conidia/ml on wheat leaves and, thus, would unlikely occur on wheat showing blast symptoms. In addition, C17-F/-R amplicon sequencing confirmed that the target sequence in IN0113 is comparable with non*Triticum* lineage isolates but differs by two SNPs, thus allowing weak primer annealing and, therefore, resulting in a late amplification signal by PCR. This nonspecific amplification could be avoided by increasing the hybridization temperature during the amplification cycle.

It was essential to evaluate the performance values of the C17 test in order to validate the method and highlight potential drawbacks. In our conditions, the test proved highly repeatable and reproducible. The limit of detection was very low, with the test consistently able to detect as few as 13 plasmid copies of the target per PCR tube (or 1 pg of genomic wheat blast DNA), and the detection of *Triticum* lineage DNA in wheat seed was successful. Our experiments confirmed the test's robustness by showing that small errors or drifts in temperature or volume would not qualitatively affect its sensitivity. However, we observed that using other brands or types of real-time PCR master mix than the ones originally used here for optimization may lead to weak cross-amplification of DNA from *Eleusine* isolate IN0113. This transferability assay confirmed that caution is required when a PCR protocol is implemented in conditions deviating from the original description (Ioos et al. 2019). This observation has recently been confirmed by additional data regarding the specificity of the MoT3 test (Gupta et al. 2019; Yasuhara-Bell et al. 2019).

The C17 diagnostic test developed here improves the detection of wheat blast to the extent of detecting *Triticum* lineage isolates that had, in previously available tests, been undetectable. On the one hand, the C17 test proved able to detect all epidemic wheat blast isolates in our possession. They consisted of 20 isolates belonging to the *Triticum* lineage tested by in silico analysis, and 30 isolates sampled on wheat with blast symptoms that were then tested by wet lab analysis. These 48 isolates had been sampled between 1988 and 2018 in five countries affected by the disease and, thus, cover a significant proportion of their genetic diversity. In our conditions, the MoT3 test (Pieck et al. 2017) failed to yield positive results for two of them (BR0032 and BR0043) by real-time PCR. On the other hand, the C17 test yielded false-positive results with DNA from three *P. oryzae* isolates not sampled on wheat; namely, CR0023, CR0057, and JP0033. We demonstrated that all three isolates were unable to induce any symptoms when inoculated by spray on a susceptible variety of wheat. Therefore, the presence of these genotypes on wheat blast showing symptoms is very unlikely and false-positive results are not expected when testing wheat tissue with the C17 test.

Prior to this study, it had been hypothesized that the isolates causing wheat blast epidemics belong to a species distinct from *P. oryzae*, and the concept of host specificity within *P. oryzae* had also been questioned (Castroagudín et al. 2016; Ceresini et al. 2019, 2018). However, further studies had refuted this hypothesis by explaining the biases that may have led to such a conclusion (Gladieux et al. 2018; Valent et al. 2019). In the present study, we chose to develop a test targeting isolates from the *Triticum* lineage of *P. oryzae* as described by Gladieux et al. (2018), assuming that only isolates belonging to the *Triticum* lineage are able to cause wheat blast epidemics. Under this assumption, the C17 test represents an efficient means to quickly assess whether any *P. oryzae* isolate sampled on wheat could lead to a wheat blast epidemic (with a positive C17 result) or, alternatively, belongs to a closely related lineage, only capable of causing opportunistic wheat infections (with a negative C17 result), such as the WBKY11 isolate in Kentucky (Farman et al. 2017). This rationale was validated because all of the isolates sampled on wheat affected by wheat blast epidemics yielded positive results with the C17 test. However, it is known that isolates belonging to other *P. oryzae* lineages can cause opportunistic infections on wheat and, if a new independent host jump were to occur with such isolates, it is very likely that these isolates would have a distinct genetic background and would remain undetectable by tests targeting markers specific to the *Triticum* lineage such as the one described here. Therefore, new wheat blast outbreaks should be thoroughly investigated.

The C17 test would be useful in studying the epidemiology and spread of wheat blast. In particular, it could be used to assess whether other members of the Poaceae family are a reservoir for *Triticum* lineage isolates. It is important to note, however, that some isolates sampled on hosts other than wheat generated false-positive results. Therefore, when testing isolates sampled on other species of Poaceae, all positive results should be confirmed using at least one other complementary test (DNA based or by pathotyping the isolate), especially in areas where wheat blast has not previously been reported. Finally, using an appropriate sampling strategy, the real-time PCR tool we developed may also be very useful for safeguarding wheat seed trade, and preventing the introduction of wheat blast disease in areas where it has been absent thus far. Wheat blast is seed borne and seed transmitted (Goulart 1988). Our preliminary experiments on artificially contaminated wheat seed suggest that the C17 test is able to detect *Triticum* lineage DNA in this matrix but, nonetheless, this result should be validated with naturally contaminated seed lots.

Literature Cited

Aman, A. 2016. "Wheat blast" threatens yield. Daily Star. https://www.thedailystar.net/backpage/wheat-blast-threatens-yield-784372

Anjos, J., Da Silva, D. B., D'Avila Charchar, M. J., and Rodrigues, G. C. 1996. Occurrence of blast fungus (*Pyricularia grisea*) on wheat and rye in the savanna region of central Brazil. Pesq. Agropec. Bras. 31:79-82.

Asuyama, H. 1963. Morphology, taxonomy, host range, and life cycle of *Piricularia oryzae*. In: The Rice Blast Disease: Proc. Symp. Int. Rice Res. Inst. John Hopkins Press. Baltimore, MD, U.S.A.

Bühlmann, A., Pothier, J. F., Tomlinson, J. A., Frey, J. E., Boonham, N., Smits, T. H. M., and Duffy, B. 2013. Genomics-informed design of loop-mediated isothermal amplification for detection of phytopathogenic *Xanthomonas arboricola* pv. *pruni* at the intraspecific level. Plant Pathol. 62:475-484.

Castroagudín, V. L., Moreira, S. I., Pereira, D. A. S., Moreira, S. S., Brunner, P. C., Maciel, J. L. N., Crous, P. W., McDonald, B. A., Alves, E., and Ceresini, P. C. 2016. *Pyricularia graminis-tritici*, a new *Pyricularia* species causing wheat blast. Persoonia Mol. Phylogeny Evol. Fungi 37:199-216.

Ceresini, P. C., Castroagudín, V. L., Rodrigues, F. Á., Rios, J. A., Aucique-Pérez, C. E., Moreira, S. I., Croll, D., Alves, E., de Carvalho, G., Maciel, J. L. N., and McDonald, B. A. 2019. Wheat blast: From its origins in South America to its emergence as a global threat. Mol. Plant Pathol. 20:155-172.

Ceresini, P. C., Castroagudín, V. L., Rodrigues, F. Á., Rios, J. A., Eduardo Aucique-Pérez, C., Moreira, S. I., Alves, E., Croll, D., and Maciel, J. L. N. 2018. Wheat blast: Past, present, and future. Annu. Rev. Phytopathol. 56: 427-456.

Cesari, S., Thilliez, G., Ribot, C., Chalvon, V., Michel, C., Jauneau, A., Rivas, S., Alaux, L., Kanzaki, H., Okuyama, Y., Morel, J.-B., Fournier, E., Tharreau, D., Terauchi, R., and Kroj, T. 2013. The rice resistance protein pair RGA4/RGA5 recognizes the *Magnaporthe oryzae* effectors AVR-Pia and AVR1-CO39 by direct binding. Plant Cell 25:1463-1481.

Couch, B. C., Fudal, I., Lebrun, M.-H., Tharreau, D., Valent, B., van Kim, P., Nottéghem, J.-L., and Kohn, L. M. 2005. Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics 170:613-630.

Cruz, C. D., and Valent, B. 2017. Wheat blast disease: Danger on the move. Trop. Plant Pathol. 42:210-222.

Farman, M., Peterson, G., Chen, L., Starnes, J., Valent, B., Bachi, P., Murdock, L., Hershman, D., Pedley, K., Fernandes, J. M., and Bavaresco, J. 2017. The *Lolium* pathotype of *Magnaporthe oryzae* recovered from a single blasted wheat plant in the United States. Plant Dis. 101:684-692.

Feau, N., Beauseigle, S., Bergeron, M.-J., Bilodeau, G. J., Birol, I., Cervantes-Arango, S., Dhillon, B., Dale, A. L., Herath, P., Jones, S. J. M., Lamarche, J., Ojeda, D. I., Sakalidis, M. L., Taylor, G., Tsui, C. K. M., Uzunovic, A., Yueh, H., Tanguay, P., and Hamelin, R. C. 2018. Genome-enhanced detection and identification (GEDI) of plant pathogens. PeerJ 6:e4392.

Fernandez, J., and Orth, K. 2018. Rise of a cereal killer: The biology of *Magnaporthe oryzae* biotrophic growth. Trends Microbiol. 26:582:597.

Firth, C., and Lipkin, W. I. 2013. The genomics of emerging pathogens. Annu. Rev. Genomics Hum. Genet. 14:281-300.

Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., and Gurr, S. J. 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484:186-194.

Franco Ortega, S., Tomlinson, J., Gilardi, G., Spadaro, D., Gullino, M. L., Garibaldi, A., and Boonham, N. 2018. Rapid detection of *Fusarium oxysporum* f. sp. *lactucae* on soil, lettuce seeds and plants using loop-mediated isothermal amplification. Plant Pathol. 67:1462-1473.

Gallet, R., Fontaine, C., Bonnot, F., Milazzo, J., Tertois, C., Adreit, H., Ravigné, V., Fournier, E., and Tharreau, D. 2016. Evolution of compatibility range in the rice-*Magnaporthe oryzae* system: An uneven distribution of R genes between rice subspecies. Phytopathology 106:348-354.

Garcez, P. P., Loiola, E. C., Madeiro da Costa, R., Higa, L. M., Trindade, P., Delvecchio, R., Nascimento, J. M., Brindeiro, R., Tanuri, A., and Rehen, S. K. 2016. Zika virus impairs growth in human neurospheres and brain organoids. Science 352:816-818.

Giraud, T., Gladieux, P., and Gavrilets, S. 2010. Linking the emergence of fungal plant diseases with ecological speciation. Trends Ecol. Evol. 25:387-395.

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., Mitchell, T., Pedley, K. F., Valent, B., Talbot, N. J., Farman, M., and Fournier, E. 2018. Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. Mol. Biol. 9:e01219-17.

Gladieux, P., Feurtey, A., Hood, M. E., Snirc, A., Clavel, J., Dutech, C., Roy, M., and Giraud, T. 2015. The population biology of fungal invasions. Mol. Ecol. 24:1969-1986.

Goulart, A. 1992. Incidência da brusone (*Pyricularia oryzae*) em diferentes cultivares de trigo (*Triticum aestivum*) em condições de campo. Fitopatol. Bras. 17:321-325.

Goulart, A. C. P. 1988. Doenças transmitidas por sementes de trigo em Mato Grosso do Sul e testes de laboratório para identificação De *Pyricularia oryzae* cav. Em sementes de trigo. Embrapa, Dourados, Brazil.

Goulart, A. C. P., Sousa, P. G., and Urashima, A. S. 2007. Danos em trigo causados pela infecção de *Pyricularia grisea* [Damages in wheat caused by infection of *Pyricularia grisea*]. Summa Phytopathol. 33:358-363.

Gupta, D. R., Reyes Avila, C. S., Win, J., Soanes, D. M., Ryder, L. S., Croll, D., Bhattacharjee, P., Hossain, M. S., Mahmud, N. U., Mehebub, M. S., Surovy, M. Z., Rahman, M. M., Talbot, N. J., Kamoun, S., and Islam, M. T. 2019. Cautionary notes on use of the MoT3 diagnostic assay for *Magnaporthe oryzae* wheat and rice blast isolates. Phytopathology 109:504-508.

Igarashi, S., Igarashi, L. C., Utiamada, C. M., Kasuma, A. H., and Lopez, R. S. 1986. *Pyricularia* sp. em trigo. Ocurrencia de *Pyricularia* sp no Estado do Parana. Fitopatol. Bras. 11:351-352.

Inoue, Y., Vy, T. T. P., Yoshida, K., Asano, H., Mitsuoka, C., Asuke, S., Anh, V. L., Cumagun, C. J. R., Chuma, I., Terauchi, R., Kato, K., Mitchell, T., Valent, B., Farman, M., and Tosa, Y. 2017. Evolution of the wheat blast fungus through functional losses in a host specificity determinant. Science 357:80-83.

Ioos, R., Aloi, F., Piškur, B., Guinet, C., Mullett, M., Berbegal, M., Bragança, H., Cacciola, S. O., Oskay, F., Comejo, C., Adamson, K., Douanla-Meli, C., Kačergius, A., Martínez-Álvarez, P., Nowakowska, J. A., Luchi, N., Vettraino, A. M., Ahumada, R., Pasquali, M., Fourie, G., Kanetis, L., Alves, A., Ghelardini, L., Dvořák, M., Sanz-Ros, A., Diez, J. J., Baskarathevan, J., and Aguayo, J. 2019. Transferability of PCR-based diagnostic protocols: An international collaborative case study assessing protocols targeting the quarantine pine pathogen *Fusarium circinatum*. Sci. Rep. 9: Article 8195.

Ioos, S., Mallet, H.-P., Leparc Goffart, I., Gauthier, V., Cardoso, T., and Herida, M. 2014. Current Zika virus epidemiology and recent epidemics. Med. Mal. Infect. 44:302-307.

Islam, M. T., Croll, D., Gladieux, P., Soanes, D. M., Persoons, A., Bhattacharjee, P., Hossain, M. S., Gupta, D. R., Rahman, M. M., Mahboob, M. G., Cook, N., Salam, M. U., Surovy, M. Z., Sancho, V. B., Maciel, J. L. N., NhaniJúnior, A., Castroagudín, V. L., de Assis Reges, J. T., Ceresini, P. C., Ravel, S., Kellner, R., Fournier, E., Tharreau, D., Lebrun, M.-H., McDonald, B. A., Stitt, T., Swan, D., Talbot, N. J., Saunders, D. G. O., Win, J., and Kamoun, S. 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. BMC Biol. 14: Article 84.

Islam, M. T., Kim, K.-H., and Choi, J. 2019. Wheat blast in Bangladesh: The current situation and future impacts. Plant Pathol. J. 35:1-10.

Kang, S., Sweigard, J. A., and Valent, B. 1995. The *PWL* host specificity gene family in the blast fungus *Magnaporthe grisea*. Mol. Plant-Microbe Interact. 8:939-948.

Kato, H., Yamamoto, M., Yamaguchi-ozaki, T., Kadouchi, H., Iwamoto, Y., Nakayashiki, H., Tosa, Y., Mayama, S., and Mori, N. 2000. Pathogenicity, mating ability and DNA restriction fragment length polymorphisms of *Pyricularia* populations isolated from *Gramineae*, *Bambusideae* and *Zingiberaceae* plants. J. Gen. Plant Pathol. 66:30-47.

Klaubauf, S., Tharreau, D., Fournier, E., Groenewald, J. Z., Crous, P. W., de Vries, R. P., and Lebrun, M.-H. 2014. Resolving the polyphyletic nature of *Pyricularia* (*Pyriculariaceae*). Stud. Mycol. 79:85-120.

Loconsole, G., Potere, O., Boscia, D., Altamura, G., Djelouah, K., Elbeaino, T., Frasheri, D., Lorusso, D., Palmisano, F., Pollastro, P., Silletti, M. R., Trisciuzzi, N., Valentini, F., Savino, V., and Saponari, M. 2014. Detection of *Xylella fastidiosa* in olive trees by molecular and serological methods. J. Plant Pathol. 96:7-14.

Maciel, J. L. N. 2011. *Magnaporthe oryzae*, the blast pathogen: Current status and options for its control. CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour. 6: 1-8.

Malaker, P. K., Barma, N. C. D., Tiwari, T. P., Collis, W. J., Duveiller, E., Singh, P. K., Joshi, A. K., Singh, R. P., Braun, H.-J., Peterson, G. L., Pedley, K. F., Farman, M. L., and Valent, B. 2016. First report of wheat blast caused by *Magnaporthe oryzae* pathotype *triticum* in Bangladesh. Plant Dis. 100: 2330.

Malik, S. A., and Khan, M. A. 1944. Parasitic fungi of the north-west Frontier Province. Indian J. Agric. Sci. 13:522-527.

Mcrae, W. 1922. Report of the imperial mycologist: Pusa Scientific Research Institute scientific reports 1921–22. Rev. Appl. Mycol. 2:258-260.

Milazzo, J., Pordel, A., Ravel, S., and Tharreau, D. 2019. First scientific report of *Pyricularia oryzae* causing gray leaf spot disease on perennial ryegrass (*Lolium perenne*) in France. Plant Dis. 103:1024.

Morse, S. S. 2001. Factors in the emergence of infectious diseases. Pages 8-26 in: Plagues and Politics: Infectious Disease and International Policy. A. T. Price-Smith, ed. Global Issues Series. Palgrave Macmillan, London, U.K.

O'Donnell, K., Sarver, A. J. B., Brandt, M., Chang, C. D., Noble-Wang, J., Park, J. B., and Sutton, A. D. 2007. Phylogenetic diversity and microsphere array-based genotyping of human pathogenic Fusaria, including isolates from the multistate contact lens-associated U.S. Keratitis outbreaks of 2005 and 2006. J. Clin. Microbiol. 45:2235-2248.

Ou, S. H. 1985. Rice Diseases. IRRI, CAB International, Kew, United Kingdom.

Picinini, E. C., and Fernandes, J. M. C. 1990. Occurrence of wheat blast (*Pyricularia oryzae*) in commercial fields in the state of Rio Grande do Sul, Brazil. Fitopatol. Bras. 15:83-84.

Pieck, M. L., Ruck, A., Farman, M. L., Peterson, G. L., Stack, J. P., Valent, B., and Pedley, K. F. 2017. Genomics-based marker discovery and diagnostic assay development for wheat blast. Plant Dis. 101:103-109.

Rush, M. C., and Carver, R. B. 1973. Ryegrass blast: A serious new disease in Louisiana. La. Agric. 16:15.

Schena, L., Nigro, F., Ippolito, A., and Gallitelli, D. 2004. Real-time quantitative PCR: A new technology to detect and study phytopathogenic and antagonistic fungi. Eur. J. Plant Pathol. 110:893-908.

Singh, N., and Kapoor, R. 2018. Quick and accurate detection of *Fusarium oxysporum* f. sp. *carthami* in host tissue and soil using conventional and real-time PCR assay. World J. Microbiol. Biotechnol. 34:175.

Te Beest, D. O., Guerber, C., and Ditmore, M. 2007. Rice Blast. Plant Health Instruct. doi:10.1094/PHI-I-2007-0313-07 https://www.apsnet.org/edcenter/disandpath/fungalasco/pdlessons/Pages/RiceBlast.aspx

Tosa, Y., Hirata, K., Tamba, H., Nakagawa, S., Chuma, I., Isobe, C., Osue, J., Urashima, A. S., Don, L. D., Kusaba, M., Nakayashiki, H., Tanaka, A., Tani, T., Mori, N., and Mayama, S. 2004. Genetic constitution and pathogenicity of *Lolium* isolates of *Magnaporthe oryzae* in comparison with host species-specific pathotypes of the blast fungus. Phytopathology 94:454-462.

Valent, B., Farman, M., Tosa, Y., Begerow, D., Fournier, E., Gladieux, P., Islam, M. T., Kamoun, S., Kemler, M., Kohn, L. M., Lebrun, M.-H., Stajich, J. E., Talbot, N. J., Terauchi, R., Tharreau, D., and Zhang, N. 2019. *Pyricularia graminis-tritici* is not the correct species name for the wheat blast fungus: Response to Ceresini et al. (MPP 20:2). Mol. Plant Pathol. 20:173-179.

Van Dam, P., De Sain, M., Ter Horst, A., Van der Gragt, M., and Rep, M. 2018. Use of comparative genomics-based markers for discrimination of host specificity in *Fusarium oxysporum*. Appl. Environ. Microbiol. 84: e01868-17.

White, T. J., Bruns, T. D., Lee, S. B., and Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 in: PCR Protocols: A Guide to Methods and Applications. M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds. Academic Press, San Diego, CA, U.S.A.

Yasuhara-Bell, J., Pedley, K. F., Farman, M., Valent, B., and Stack, J. P. 2018. Specific detection of the wheat blast pathogen (*Magnaporthe oryzae* Triticum) by loop-mediated isothermal amplification. Plant Dis. 102:2550-2559.

Yasuhara-Bell, J., Pieck, M. L., Ruck, A., Farman, M., Peterson, G. L., Stack, J. P., Valent, B., and Pedley, K. F. 2019. A response to Gupta et al. (2019) regarding the MoT3 wheat blast diagnostic assay. Phytopathology 109:509-511.

SUPPLEMENTARY MATERIAL

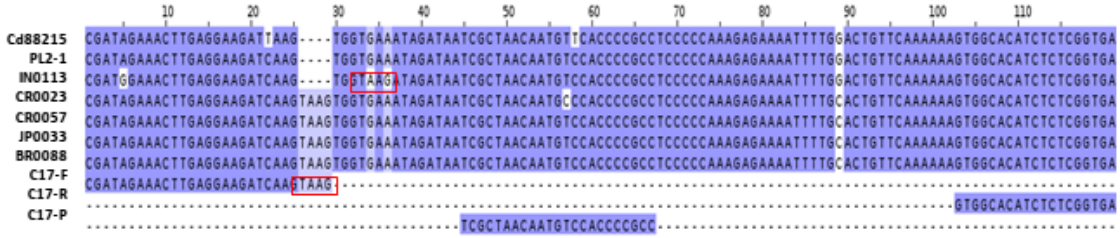


Fig. S1. Genomic region sequence targeted by primer set 17 (forward primer: C17-F, reverse primer: C17-R, and probe: C17-P). Represented sequences are: wheat-borne isolate BR0088; cross-amplified isolates JP0033, CR0023, and CR0057; isolate IN0113 slightly amplified with lower hybridization temperature; unamplified non-wheat-borne isolates Cd88215 and PL2-1; and the C17 primer set's primers and probe sequence. Regions were amplified using the external primers seqF_C17 and seqR_C17. Red boxes indicate the sequences shared between non-wheat-borne isolate IN0113 and the 3' end of the forward primer C17-F.

Fig. S2. Compatibility assessment between Pyricularia oryzae isolates of wheat and ryegrass
Symptoms observed 1 week after inoculation of a conidial suspension on 2-week-old seedlings. The inoculation was done either by spraying a 20,000 conidia/ml suspension on the whole plant or by applying a 50,000 conidia/ml suspension using a brush in a delimited leaf region. These inoculations confirmed the compatibility between the wheat variety Thésée and the wheat-borne isolates tested (BR0032 and BL0028) and, on the contrary, the lack of compatibility with the non-wheat-borne isolates [CR0023 (sampled on *Echinochloa crus-galli*), CR0057 (*Lolium sp*.), JP0033 (*Eriochloa villosa*), and AG0055 (*Bromus spp*.)].

Table S1. Genome assemblies of *Pyricularia oryzae*, *P. grisea* and *P. pennisetigena* used in this study.

Isolate ID	Synonym	Species	Host	Lineage	Year	Locality	NCBI accession number
BR29	BR0029	Pyricularia grisae	Digitaria sanguinalis	ND	1989	Brazil	
Dig41		Pyricularia grisae	Digitaria sanguinalis	ND	ND	Hyogo, Japan	
DsLIZ		Pyricularia grisae	Digitaria sanguinalis	ND	2000	Lexington, KY, USA	SAMN08009550
VO107		Pyricularia grisae	Digitaria sanguinalis	ND	1981	Texas, USA	SAMN08009577
Bm88324		Pyricularia oryzae	Brachiaria mutica	Brachiaria1	1988	The Philippines	SAMN08009544
Bd8401		Pyricularia oryzae	Brachiaria distachya	Brachiaria2	1984	The Philippines	SAMN08009543
B51		Pyricularia oryzae	Eleusine indica	Eleusine1	2012	Quirusillas, Bolivia	SAMN08009542
BR62		Pyricularia oryzae	Eleusine indica	Eleusine1	1991	Brazil	
CD156	CD0156	Pyricularia oryzae	Eleusine indica	Eleusine1	1989	Ferkessedougou, Ivory Coast	
EI9604		Pyricularia oryzae	Eleusine indica	Eleusine1	1996	Fujian, China	
PH42		Pyricularia oryzae	Eleusine coracana	Eleusine1	1983	The Philippines	SAMN08009570
EI9411		Pyricularia oryzae	Eleusine indica	Eleusine2	1990	Fujian, China	
G22	WGG-FA40	Pyricularia oryzae	Eleusine coracana	Eleusine2	1976	Japan	SAMN08009554
Z2-1		Pyricularia oryzae	Eleusine coracana	Eleusine2	1977	Kagawa, Japan	
G17	K76-79	Pyricularia oryzae	Eragrostis curvula	Eragrostis	1976	Japan	SAMN08009553
Br58		Pyricularia oryzae	Avena sativa	Lolium	1990	Parana, Brazil	
CHRF		Pyricularia oryzae	Lolium perenne	Lolium	1996	Silver Spring, MD, USA	SAMN08009548
CHW		Pyricularia oryzae	Lolium perenne	Lolium	1996	Annapolis, MD, USA	SAMN08009549
FH		Pyricularia oryzae	Lolium perenne	Lolium	1997	Hagerstown, MD, USA	SAMN08009551
GG11		Pyricularia oryzae	Lolium perenne	Lolium	1997	Lexington, KY, USA	SAMN08009555
HO		Pyricularia oryzae	Lolium perenne	Lolium	1996	Richmond, PA, USA	SAMN08009558
LpKY97	LpKY97-1	Pyricularia oryzae	Lolium perenne	Lolium	1997	Lexington, KY, USA	SAMN08009564
P28	P-0028	Pyricularia oryzae	Bromus tectorum	Lolium	2014	Paraguay	SAMN05864041
Pg1213-22		Pyricularia oryzae	Festuca arundinaceum	Lolium	1999/2000	GA	SAMN08009569
PgKY	PgKY4OV2.1	Pyricularia oryzae	Lolium perenne	Lolium	2000	Lexington, KY, USA	
PGPA	PgPA18C-02, PgPA	Pyricularia oryzae	Lolium perenne	Lolium	1998	Pennsylvania, USA	
PL2-1		Pyricularia oryzae	Lolium multiflorum	Lolium	2002	Pulaski Co., KY, USA	SAMN08009571
PL3-1		Pyricularia oryzae	Lolium multiflorum	Lolium	2002	Pulaski Co., KY, USA	SAMN08009572
PY5010	PY05010	Pyricularia oryzae	Triticum aestivum	Lolium	2005	Londrina, Brazil	
PY86-1	PY86.1	Pyricularia oryzae	Triticum aestivum	Lolium	2008	Cascavel, Brazil	
TF05-1		Pyricularia oryzae	Festuca arundinaceum	Lolium	2005	Lexington, KY, USA	SAMN08009576
WBKY11	WBKY11-15	Pyricularia oryzae	Triticum aestivum	Lolium	2011	Lexington, KY, USA	SAMN08009578
87-120		Pyricularia oryzae	Oryza sativa	Oryza	ND		PQBK00000000
FR13	FR0013	Pyricularia oryzae	Oryza sativa	Oryza	1988	France	
GY11	GY0011, Guy11	Pyricularia oryzae	Oryza sativa	Oryza	1988	French Guiana	
IA1	ARB114	Pyricularia oryzae	Oryza sativa	Oryza	2009	Arkansas, USA	SAMN08009559
IB33		Pyricularia oryzae	Oryza sativa	Oryza	ND	Texas, USA	SAMN08009560
IB49	ZN61	Pyricularia oryzae	Oryza sativa	Oryza	1992	Arkansas, USA	SAMN08009561
IC17	ZN57	Pyricularia oryzae	Oryza sativa	Oryza	1992	Arkansas, USA	SAMN08009562
IE1K	TM2	Pyricularia oryzae	Oryza sativa	Oryza	2003	Arkansas, USA	SAMN08009563
INA168	Ina168	Pyricularia oryzae	Oryza sativa	Oryza	1958	Aichi, Japan	
KEN53-33	Ken53-33	Pyricularia oryzae	Oryza sativa	Oryza	1953	Aichi, Japan	
ML33		Pyricularia oryzae	Oryza sativa	Oryza	1995	Mali	SAMN08009565
P131		Pyricularia oryzae	Oryza sativa	Oryza	ND	Japan	
P-2	P2	Pyricularia oryzae	Oryza sativa	Oryza	1948	Aichi, Japan	
PH0014-rn	PH0014, PH14	Pyricularia oryzae	Oryza sativa	Oryza	ND	The Philippines	
TH0012-rn	TH0012, TH12	Pyricularia oryzae	Hordeum vulgare	Oryza	ND	Thailand	
TH0016	TH16	Pyricularia oryzae	Hordeum vulgare	Oryza	ND	Thailand	
TH3		Pyricularia oryzae	Oryza sativa	Oryza	ND	Thailand	
Y34		Pyricularia oryzae	Oryza sativa	Oryza	1982	Yunnan, China	
Arcadia		Pyricularia oryzae	Setaria viridis	Setaria	1998	Lexington, KY, USA	SAMN08009540
GFSI1-7-2	GFSI	Pyricularia oryzae	Setaria italica	Setaria	1977	Gifu, Japan	
GrF52		Pyricularia oryzae	Setaria viridis	Setaria	2001	Lexington, KY, USA	SAMN08009556
KANSV1-4-	KNSV	Pyricularia oryzae	Setaria viridis	Setaria	1975	Kanagawa, Japan	
SA05-43		Pyricularia oryzae	Setaria viridis	Setaria	2005	Nagasaki, Japan	
Sv9610		Pyricularia oryzae	Setaria viridis	Setaria	1996	Zhejian, China	
Sv9623		Pyricularia oryzae	Setaria viridis	Setaria	1996	Zhejian, China	
US0071	US71	Pyricularia oryzae	Setaria spp.	Setaria	ND	USA	
SSFL02		Pyricularia oryzae	Stenotaphrum secundatum	Stenotaphrum	2002	Disneyworld, FL, USA	SAMN08009573
SSFL14-3		Pyricularia oryzae	Stenotaphrum secundatum	Stenotaphrum	2014	New Smyrna, FL, USA	SAMN08009574
B2		Pyricularia oryzae	Triticum aestivum	Triticum	2011	Bolivia	SAMN05580113
B71		Pyricularia oryzae	Triticum aestivum	Triticum	2012	Bolivia	SAMN04942725
BdBar	BdBar16-1	Pyricularia oryzae	Triticum aestivum	Triticum	2016	Barisal, Bangladesh	SAMN04940126
BdJes	BdJes16-1	Pyricularia oryzae	Triticum aestivum	Triticum	2016	Jessore, Bangladesh	SAMN04942531
BdMeh	BdMeh16-1	Pyricularia oryzae	Triticum aestivum	Triticum	2016	Mehepur, Bangladesh	SAMN04942534
BR0032	BR32	Pyricularia oryzae	Triticum aestivum	Triticum	1991	Brazil	
Br130		Pyricularia oryzae	Triticum aestivum	Triticum	1990	Mato Grosso do Sul, Brazil	SAMN08009547
Br48		Pyricularia oryzae	Triticum aestivum	Triticum	1990	Mato Grosso do Sul, Brazil	
Br7		Pyricularia oryzae	Triticum aestivum	Triticum	1990	Parana, Brazil	SAMN08009545
Br80		Pyricularia oryzae	Triticum aestivum	Triticum	1991	Brazil	SAMN08009546
P29	P-0029	Pyricularia oryzae	Bromus tectorum	Triticum	2014	Paraguay	SAMN05898532
P3		Pyricularia oryzae	Triticum durum	Triticum	2012	Canindeyu, Paraguay	SAMN08009568
PY0925		Pyricularia oryzae	Triticum aestivum	Triticum	2009	Predizes, Brazil	
PY36-1	PY36.1	Pyricularia oryzae	Triticum aestivum	Triticum	2007	Brasilia, Brazil	
PY5003	PY05003	Pyricularia oryzae	Triticum aestivum	Triticum	2005	Londrina, Brazil	
PY5033	PY05033	Pyricularia oryzae	Triticum aestivum	Triticum	2005	Londrina, Brazil	
PY6017	PY06017	Pyricularia oryzae	Triticum aestivum	Triticum	2006	Coromandel, Brazil	
PY6045	PY06045	Pyricularia oryzae	Triticum aestivum	Triticum	2006	Goiania, Brazil	
T25		Pyricularia oryzae	Triticum aestivum	Triticum	1988	Parana, Brazil	SAMN08009575
WHTQ		Pyricularia oryzae	Triticum aestivum	Triticum	ND	Brazil	SAMN08009580
PM1		Pyricularia pennisetigena	Pennisetum americanum	ND	1990	Georgia, USA	PQBJ00000000

Adapted from Gladieux and al., 2018

ND: no data.

Table S2. Characteristics of the DNA samples used in this study.

Isolate	Alternative name	Species	Host	Host-specific lineage*	Year of isolation	DNA origin
Br58		Pyricularia oryzae	Avena sp.	Lolium	1990	Iwate Biotechnology Research Center
PH0075	G159	Pyricularia oryzae	Brachiaria mutica		1989	UMR BGPI
AG0054	706	Pyricularia oryzae	Bromus sp.		2002	UMR BGPI
AG0055	707	Pyricularia oryzae	Bromus sp.		2002	UMR BGPI
P28		Pyricularia oryzae	Bromus sp.	Lolium	2014	Department of Plant Pathology, University of Kentucky
P29		Pyricularia oryzae	Bromus sp.	Triticum	2014	Department of Plant Pathology, University of Kentucky
AG0061	812	Pyricularia oryzae	Bromus unioloides		2003	UMR BGPI
US0066	BG11-1-3	Pyricularia oryzae	Cenchrus ciliaris		1995	UMR BGPI
BR0030		Pyricularia oryzae	Cenchrus echinatus		1989	UMR BGPI
GR0001	CT4	Pyricularia oryzae	Ctenanthe oppenheimiana		1998	UMR BGPI
EG0028	27	Pyricularia oryzae	Cyperus rotundus			UMR BGPI
IS0001	PyH1,CBS 665.79	Pyricularia oryzae	Cyperus rotundus		1979	UMR BGPI
PH0052	CRA8401	Pyricularia oryzae	Cyperus rotundus		1990	UMR BGPI
PH0053	CR88383	Pyricularia oryzae	Cyperus rotundus		1990	UMR BGPI
CD0143		**Pyricularia grisea**	Digitaria exilis		1989	UMR BGPI
BR0029		**Pyricularia grisea**	Digitaria sanguinalis		1989	UMR BGPI
CR0023	KAP-9	Pyricularia oryzae	Echinochloa crus-galli			UMR BGPI
Cd8821 5	PH0051	Pyricularia oryzae	Cynodon dactylon		1988	Department of Plant Pathology, University of Kentucky
IR0102	42EAZ	Pyricularia oryzae	Echinochloa sp.		2016	UMR BGPI
PH0078	Ec-A8401,G164	Pyricularia oryzae	Echinochloa sp.		1989	UMR BGPI
BR0062		Pyricularia oryzae	Eleusine indica		1990	UMR BGPI
IN0113	VII-953-1	Pyricularia oryzae	Eleusine sp.		1992	UMR BGPI
RW0043	G199	Pyricularia oryzae	Eleusine coracana			UMR BGPI
BR0070		Pyricularia oryzae	Eragrostis sp.		1991	UMR BGPI
JP0028	K76-79,G-17,US27	Pyricularia oryzae	Eragrostis curvula		1976	UMR BGPI
JP0033	NI859	Pyricularia oryzae	Eriochloa villosa			UMR BGPI
CR0029	KAP-20	Pyricularia oryzae	Festuca elalior			UMR BGPI
Pg1213-22		Pyricularia oryzae	Festuca sp.	Lolium	1999/2000	Department of Plant Pathology, University of Kentucky
TF05-1		Pyricularia oryzae	Festuca sp.	Lolium	2005	Department of Plant Pathology, University of Kentucky
JP0047	M80-25	Pyricularia oryzae	Hordeum vulgare		1980	UMR BGPI
JP0048	M80-28	Pyricularia oryzae	Hordeum vulgare		1980	UMR BGPI
KN0001		Pyricularia oryzae	Hordeum vulgare		1994	UMR BGPI
KN0006		Pyricularia oryzae	Hordeum vulgare		1994	UMR BGPI
VT0032	DR51-2-3	Pyricularia oryzae	Leersia hexandra		2002	UMR BGPI
Lc8401	PH0060	Pyricularia oryzae	Leptochloa chimensis		1984	Department of Plant Pathology, University of Kentucky
US0077	330	Pyricularia oryzae	Lolium perenne			UMR BGPI
US0078	365	Pyricularia oryzae	Lolium perenne			UMR BGPI
AG0062	822	Pyricularia oryzae	Lolium sp.		2004	UMR BGPI
AG0063	823	Pyricularia oryzae	Lolium sp.		2004	UMR BGPI
AG0064	824	Pyricularia oryzae	Lolium sp.		2004	UMR BGPI
CHRF		Pyricularia oryzae	Lolium sp.	Lolium	1996	Department of Plant Pathology, University of Kentucky
CHW		Pyricularia oryzae	Lolium sp.	Lolium	1996	Department of Plant Pathology, University of Kentucky
CR0026	KAP-23	Pyricularia oryzae	Lolium sp.		1991	UMR BGPI
CR0057	W95-11	Pyricularia oryzae	Lolium sp.			UMR BGPI
FH		Pyricularia oryzae	Lolium sp.	Lolium	1997	Department of Plant Pathology, University of Kentucky
FR1069		Pyricularia oryzae	Lolium sp.		2017	UMR BGPI
GG11		Pyricularia oryzae	Lolium sp.	Lolium	1997	Department of Plant Pathology, University of Kentucky
HO		Pyricularia oryzae	Lolium sp.	Lolium	1996	Department of Plant Pathology, University of Kentucky
LpKY97	LpKY97-1	Pyricularia oryzae	Lolium sp.	Lolium	1997	Department of Plant Pathology, University of Kentucky
PL2-1		Pyricularia oryzae	Lolium sp.		2002	Department of Plant Pathology, University of Kentucky
PL3-1		Pyricularia oryzae	Lolium sp.		2002	Department of Plant Pathology, University of Kentucky
AG0132	89	Pyricularia oryzae	Oryza sativa		2001	UMR BGPI
BR0019		Pyricularia oryzae	Oryza sativa		1986	UMR BGPI
CH0333		Pyricularia oryzae	Oryza sativa		1998	UMR BGPI
CH1120		Pyricularia oryzae	Oryza sativa		2009	UMR BGPI
GY0011	GUY11	Pyricularia oryzae	Oryza sativa		1978	UMR BGPI
IN0082	45b-1-1	Pyricularia oryzae	Oryza sativa		1992	UMR BGPI
IN0115	ML2	Pyricularia oryzae	Oryza sativa		1997	UMR BGPI

FR0013		Pyricularia oryzae	Oryza sp.		1988	UMR BGPI
JP0031	NI992	Pyricularia oryzae	Panicum coloratum		1991	UMR BGPI
JP0030	NI885	Pyricularia oryzae	Panicum bisulcatum		1991	UMR BGPI
IN0005	IN77-33-1-1	Pyricularia oryzae	Panicum maximun			UMR BGPI
CR0021	KAP-1	Pyricularia oryzae	Panicum miliaceum			UMR BGPI
IN0003	IN77-28-1-1	Pyricularia oryzae	Panicum repens			UMR BGPI
Pr8202		Pyricularia oryzae	Panicum repens			Department of Plant Pathology, University of Kentucky
Pd88413		Pyricularia oryzae	Paspalum distichum		1988	Department of Plant Pathology, University of Kentucky
PH0062	PD8824	Pyricularia oryzae	Paspalum distichum		1990	UMR BGPI
PH0097	G-65	Pyricularia oryzae	Paspalum paspaloides		1983	UMR BGPI
ML0031		**Pyricularia pennisetigena**	Pennisetum sp.		1990	UMR BGPI
BF0017		**Pyricularia pennisetigena**	Pennisetum typhoides		1990	UMR BGPI
CR0031	KAP-2	Pyricularia oryzae	Setaria italica		1991	UMR BGPI
IN0022	HC3-2	Pyricularia oryzae	Setaria sp.		1992	UMR BGPI
IN0023	HC5-3A	Pyricularia oryzae	Setaria sp.		1992	UMR BGPI
IN0108	VII-765-1	Pyricularia oryzae	Setaria sp.		1992	UMR BGPI
US0064	G-188	Pyricularia oryzae	Setaria sp.		1991	UMR BGPI
CR0030	KAP-11	Pyricularia oryzae	Setaria viridis		1991	UMR BGPI
Pg1054		Pyricularia oryzae	Stenotaphrum secundatum		2000	Department of Plant Pathology, University of Kentucky
PR0069		Pyricularia oryzae	Stenotaphrum secundatum		1992	UMR BGPI
US0084	SAG00T12	Pyricularia oryzae	Stenotaphrum secundatum		2000	UMR BGPI
AG0065	843	Pyricularia oryzae	Stenotaphrum sp.		2007	UMR BGPI
AG0103	ACA8	Pyricularia oryzae	Triticum aestivum		2016	UMR BGPI
BL0017		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0018		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0020		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0023		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0028		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0037		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0044		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0046		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BL0063		Pyricularia oryzae	Triticum aestivum		2010	UMR BGPI
BR0123	PR01-95	Pyricularia oryzae	Triticum aestivum		1998	UMR BGPI
BL0092		Pyricularia oryzae	Triticum sp.		2017	UMR BGPI
BL0093		Pyricularia oryzae	Triticum sp.		2017	UMR BGPI
BR0031		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0032		Pyricularia oryzae	Triticum sp.	Triticum	1991	UMR BGPI
BR0034		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0036		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0039		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0040		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0041		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0043		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0045		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0047		Pyricularia oryzae	Triticum sp.		1989	UMR BGPI
BR0080	G157,Kmetz N°22	Pyricularia oryzae	Triticum sp.			UMR BGPI
BR0086	BR3	Pyricularia oryzae	Triticum sp.		1990	UMR BGPI
BR0087	BR8	Pyricularia oryzae	Triticum sp.		1990	UMR BGPI
BR0088	BR52	Pyricularia oryzae	Triticum sp.		1990	UMR BGPI
BTGP16		Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory (isolate provided by T. Islam from BSMRAU)
BTJP4-1		Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory (isolate provided by T. Islam from BSMRAU)
BTMP13-1		Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory (isolate provided by T. Islam from BSMRAU)
GN0001	GA1	Pyricularia oryzae	Zea mays		1985	UMR BGPI
IR0013	ZG1	Pyricularia oryzae	Zea mays		2012	UMR BGPI
IR0015	ZG1-1-2	Pyricularia oryzae	Zea mays		2012	UMR BGPI
IR0095	22azga-3	Pyricularia oryzae	Zea mays		2016	UMR BGPI

* According to Gladieux et al. (2018)

Table S3. Identified genetic variants specific to the Triticum lineage.

	LOCATION IN BR0032 GENOME					
	BR0032 scaffold	**fragment start**	**fragment stop**	**position**	***Triticum* allele**	**non-*Triticum* allele**
1	scaffold00010	1275500	1276000	487	['G']	['T', '-']
2	scaffold00010	1275750	1276250	195	['G']	['T', '-']
3	scaffold00015	348000	348500	188	['T']	['-', 'A']
4	scaffold00015	350750	351250	291	['G']	['T', '-']
5	scaffold00015	351000	351500	10	['G']	['T', '-']
6	scaffold00015	355000	355500	269	['A']	['T', '-']
7	scaffold00015	355250	355750	429	['T']	['G', '-']
8	scaffold00015	355500	356000	152	['T']	['G', '-']
9	scaffold00015	373000	373500	25	['G']	['-', 'A']
10	scaffold00015	373000	373500	43	['T']	['-', 'G']
11	scaffold00015	373000	373500	48	['C']	['-', 'A']
12	scaffold00015	373000	373500	66	['C']	['-', 'A']
13	scaffold00015	373000	373500	101	['C']	['-']
14	scaffold00015	373000	373500	134	['C']	['-', 'T']
15	scaffold00015	373000	373500	140	['G']	['-', 'C']
16	scaffold00015	373000	373500	141	['A']	['-', 'T']
17	scaffold00015	373000	373500	148	['G']	['-', 'A']
18	scaffold00015	373000	373500	157	['T']	['-', 'C']
19	scaffold00015	373000	373500	166	['G']	['-', 'C']
20	scaffold00015	373000	373500	167	['G']	['-', 'A']
21	scaffold00015	373000	373500	182	['G']	['-', 'A']
22	scaffold00015	373000	373500	184	['A']	['-', 'T']
23	scaffold00015	373000	373500	189	['A']	['-', 'C']
24	scaffold00015	373000	373500	196	['T']	['-', 'C']
25	scaffold00015	373000	373500	208	['G']	['-', 'A']
26	scaffold00015	373000	373500	238	['C']	['-', 'A']
27	scaffold00015	373000	373500	241	['C']	['-', 'T']
28	scaffold00015	373000	373500	257	['C']	['-', 'T']
29	scaffold00015	373000	373500	258	['T']	['-', 'G']
30	scaffold00015	373000	373500	261	['A']	['-', 'G']
31	scaffold00015	373000	373500	266	['C']	['-', 'T']
32	scaffold00015	373000	373500	269	['T']	['-', 'C']
33	scaffold00015	373000	373500	287	['C']	['-', 'T']
34	scaffold00015	373000	373500	292	['T']	['-', 'G']
35	scaffold00015	373000	373500	351	['A']	['-', 'G']
36	scaffold00015	373250	373750	2	['C']	['-']
37	scaffold00015	373250	373750	3	['T']	['-']
38	scaffold00015	373250	373750	6	['A']	['-', 'G']
39	scaffold00015	373250	373750	11	['C']	['-', 'T']
40	scaffold00015	373250	373750	14	['T']	['-', 'C']
41	scaffold00015	373250	373750	32	['C']	['-', 'T']
42	scaffold00015	373250	373750	37	['T']	['-', 'G']
43	scaffold00015	373250	373750	51	['T']	['-', 'C']
44	scaffold00015	373250	373750	55	['C']	['-', 'G']
45	scaffold00015	373250	373750	95	['A']	['-', 'G']
46	scaffold00015	373500	374000	408	['G']	['-', 'T']
47	scaffold00015	373750	374250	146	['G']	['-', 'T']
48	scaffold00015	373750	374250	425	['A']	['-', 'G']
49	scaffold00015	373750	374250	466	['G']	['-']
50	scaffold00015	373750	374250	467	['T']	['-']
51	scaffold00015	373750	374250	468	['A']	['-']
52	scaffold00015	373750	374250	469	['C']	['-']
53	scaffold00015	373750	374250	470	['A']	['-']
54	scaffold00015	373750	374250	471	['G']	['-']
55	scaffold00015	373750	374250	472	['C']	['-']
56	scaffold00015	373750	374250	473	['G']	['-']
57	scaffold00015	373750	374250	474	['T']	['-']
58	scaffold00015	373750	374250	475	['G']	['-']
59	scaffold00015	373750	374250	476	['T']	['-']
60	scaffold00015	373750	374250	477	['G']	['-']
61	scaffold00015	373750	374250	478	['T']	['-']
62	scaffold00015	373750	374250	479	['A']	['-']
63	scaffold00015	373750	374250	480	['C']	['-']
64	scaffold00015	373750	374250	481	['A']	['-']
65	scaffold00015	373750	374250	482	['C']	['-']
66	scaffold00015	373750	374250	483	['C']	['-']
67	scaffold00015	373750	374250	484	['T']	['-']
68	scaffold00015	373750	374250	485	['G']	['-']

69	scaffold00015	373750	374250	486	['G']	['-']
70	scaffold00015	373750	374250	487	['G']	['-']
71	scaffold00015	373750	374250	489	['G']	['-']
72	scaffold00015	373750	374250	490	['A']	['-']
73	scaffold00015	373750	374250	491	['G']	['-']
74	scaffold00015	373750	374250	492	['G']	['-']
75	scaffold00015	373750	374250	493	['T']	['-']
76	scaffold00015	373750	374250	494	['G']	['-']
77	scaffold00015	373750	374250	495	['T']	['-']
78	scaffold00015	373750	374250	496	['C']	['-']
79	scaffold00015	373750	374250	497	['C']	['-']
80	scaffold00015	373750	374250	498	['T']	['-']
81	scaffold00015	373750	374250	499	['T']	['-']
82	scaffold00015	373750	374250	500	['T']	['-']
83	scaffold00015	373750	374250	501	['C']	['-']
84	scaffold00015	373750	374250	502	['C']	['-']
85	scaffold00015	373750	374250	503	['A']	['-']
86	scaffold00015	373750	374250	504	['G']	['-']
87	scaffold00015	373750	374250	505	['T']	['-']
88	scaffold00015	373750	374250	506	['C']	['-']
89	scaffold00015	373750	374250	507	['C']	['-']
90	scaffold00015	373750	374250	508	['C']	['-']
91	scaffold00015	374000	374500	170	['A']	['G', '-']
92	scaffold00015	374000	374500	271	['C']	['T', '-']
93	scaffold00015	374000	374500	440	['C']	['-', 'A']
94	scaffold00015	376250	376750	308	['G']	['A', '-']
95	scaffold00015	376500	377000	53	['G']	['-', 'A']
96	scaffold00015	376500	377000	336	['G']	['-', 'A']
97	scaffold00015	376500	377000	381	['A']	['-', 'G']
98	scaffold00015	376750	377250	80	['G']	['A', '-']
99	scaffold00015	376750	377250	125	['A']	['G', '-']
100	scaffold00015	377500	378000	221	['A']	['-', 'G']
101	scaffold00015	381250	381750	298	['T']	['-']
102	scaffold00015	381250	381750	497	['A']	['T', '-', 'G']
103	scaffold00015	381500	382000	242	['A']	['T', '-', 'G']
104	scaffold00015	381750	382250	361	['C']	['-', 'T']
105	scaffold00015	382000	382500	106	['C']	['-', 'T']
106	scaffold00015	384750	385250	445	['C']	['G', 'T', '-']
107	scaffold00015	384750	385250	479	['A']	['C', '-', 'G']
108	scaffold00015	385000	385500	187	['C']	['T', 'G', '-']
109	scaffold00015	385000	385500	221	['A']	['-', 'C', 'G']
110	scaffold00015	386500	387000	333	['A']	['G', '-']
111	scaffold00015	386750	387250	59	['A']	['G', '-']
112	scaffold00015	387000	387500	391	['C']	['T', '-']
113	scaffold00015	387250	387750	103	['C']	['T', '-']
114	scaffold00015	393500	394000	493	['A']	['-', 'T', 'G']
115	scaffold00015	394250	394750	262	['T']	['C', '-']
116	scaffold00015	394250	394750	302	['C']	['T', '-']
117	scaffold00015	394250	394750	369	['A']	['G', '-']
118	scaffold00015	394250	394750	476	['G']	['T', '-']
119	scaffold00015	394500	395000	189	['G']	['-', 'T']
120	scaffold00015	394500	395000	229	['G']	['-', 'C']
121	scaffold00015	395000	395500	418	['C']	['G', '-']
122	scaffold00015	395250	395750	143	['C']	['-', 'G']
123	scaffold00015	396000	396500	306	['T']	['G', '-', 'C']
124	scaffold00015	396250	396750	41	['T']	['G', 'C', '-']
125	scaffold00015	396750	397250	407	['A']	['-']
126	scaffold00015	397000	397500	148	['A']	['-']
127	scaffold00015	397000	397500	420	['G']	['-']
128	scaffold00015	397250	397750	114	['G']	['-']
129	scaffold00015	399500	400000	481	['A']	['G', '-']
130	scaffold00015	401250	401750	324	['A']	['-', 'G']
131	scaffold00015	401250	401750	325	['T']	['-', 'G']
132	scaffold00015	401250	401750	344	['C']	['-', 'A']
133	scaffold00015	401500	402000	87	['C']	['A', '-']
134	scaffold00015	402250	402750	326	['T']	['-']
135	scaffold00015	402250	402750	327	['G']	['-']
136	scaffold00015	402250	402750	480	['A']	['-', 'G']
137	scaffold00015	402500	403000	211	['A']	['G', '-']
138	scaffold00015	402500	403000	501	['G']	['-', 'A']
139	scaffold00015	402500	403000	503	['C']	['-']
140	scaffold00015	402500	403000	504	['A']	['-']
141	scaffold00015	402500	403000	505	['A']	['-']

142	scaffold00015	402750	403250	241	['A']	['G', '-']
143	scaffold00015	402750	403250	242	['A']	['C', '-']
144	scaffold00015	402750	403250	402	['C']	['A', 'G', '-']
145	scaffold00015	402750	403250	499	['T']	['C', '-']
146	scaffold00015	403750	404250	523	['G']	['-', 'T', 'A']
147	scaffold00015	404750	405250	322	['C']	['-', 'G', 'A']
148	scaffold00015	404750	405250	325	['G']	['-', 'C', 'A']
149	scaffold00015	404750	405250	338	['C']	['T', '-']
150	scaffold00015	412250	412750	236	['G']	['C', 'A', '-']
151	scaffold00015	412250	412750	340	['T']	['-', 'A']
152	scaffold00015	412250	412750	368	['G']	['-', 'T']
153	scaffold00015	426500	427000	447	['G']	['-', 'A']
154	scaffold00015	426500	427000	454	['A']	['-', 'G']
155	scaffold00015	426750	427250	187	['G']	['A', '-']
156	scaffold00015	426750	427250	261	['C']	['-', 'A']
157	scaffold00015	427000	427500	196	['A']	['-', 'G']
158	scaffold00015	427000	427500	457	['T']	['C', '-']
159	scaffold00015	427250	427750	203	['T']	['C', '-']
160	scaffold00015	428000	428500	72	['T']	['G', '-']
161	scaffold00015	428000	428500	208	['G']	['A', '-']
162	scaffold00015	428000	428500	260	['C']	['-']
163	scaffold00015	428000	428500	261	['T']	['-']
164	scaffold00015	428000	428500	261	['C']	['-']
165	scaffold00015	428000	428500	262	['A']	['-']
166	scaffold00015	428000	428500	263	['C']	['-']
167	scaffold00015	428000	428500	264	['G']	['-']
168	scaffold00015	428000	428500	265	['T']	['-']
169	scaffold00015	428000	428500	266	['A']	['-']
170	scaffold00015	428000	428500	267	['T']	['-']
171	scaffold00015	428000	428500	268	['G']	['-']
172	scaffold00015	428000	428500	269	['T']	['-']
173	scaffold00015	428000	428500	271	['C']	['-', 'T']
174	scaffold00015	428250	428750	2	['C']	['-']
175	scaffold00015	428500	429000	136	['T']	['-', 'C']
176	scaffold00015	430750	431250	219	['A']	['-', 'G']
177	scaffold00015	430750	431250	413	['C']	['-', 'A', 'G']
178	scaffold00015	431000	431500	136	['C']	['G', '-']
179	scaffold00015	443500	444000	346	['A']	['T', 'C', '-']
180	scaffold00015	443750	444250	258	['T']	['G', '-']
181	scaffold00015	450250	450750	459	['T']	['C', '-']
182	scaffold00015	451250	451750	379	['G']	['A', '-']
183	scaffold00015	451500	452000	117	['G']	['A', '-']
184	scaffold00015	454000	454500	277	['C']	['T', '-']
185	scaffold00015	457750	458250	307	['A']	['T', '-']
186	scaffold00015	458000	458500	30	['A']	['T', '-']
187	scaffold00015	463250	463750	168	['A']	['C', '-']
188	scaffold00015	463250	463750	172	['G']	['T', 'C', '-']
189	scaffold00015	465500	466000	404	['A']	['T', '-']
190	scaffold00015	465750	466250	149	['A']	['T', '-']
191	scaffold00015	470250	470750	412	['T']	['-', 'G']
192	scaffold00015	470500	471000	156	['T']	['G', '-']
193	scaffold00015	471500	472000	278	['A']	['-', 'C']
194	scaffold00015	472250	472750	2	['G']	['-']
195	scaffold00015	472250	472750	3	['A']	['-']
196	scaffold00015	472250	472750	4	['C']	['-']
197	scaffold00015	472250	472750	5	['C']	['-']
198	scaffold00015	472250	472750	6	['T']	['-']
199	scaffold00015	472250	472750	7	['G']	['-']
200	scaffold00015	472250	472750	8	['C']	['-']
201	scaffold00015	472250	472750	9	['C']	['-']
202	scaffold00015	472250	472750	10	['A']	['-']
203	scaffold00015	472250	472750	11	['C']	['-']
204	scaffold00015	472250	472750	12	['T']	['-']
205	scaffold00015	472250	472750	13	['G']	['-']
206	scaffold00015	472250	472750	14	['C']	['-']
207	scaffold00015	472250	472750	15	['G']	['-']
208	scaffold00015	472250	472750	16	['G']	['-']
209	scaffold00015	472250	472750	17	['A']	['-']
210	scaffold00015	472250	472750	18	['C']	['-']
211	scaffold00015	472250	472750	19	['C']	['-']
212	scaffold00015	472250	472750	20	['G']	['-']
213	scaffold00015	472250	472750	21	['C']	['-']
214	scaffold00015	472250	472750	22	['A']	['-']

215	scaffold00015	472250	472750	23	['G']	['-']
216	scaffold00015	472250	472750	24	['C']	['-']
217	scaffold00015	472250	472750	25	['T']	['-']
218	scaffold00015	472250	472750	26	['C']	['-']
219	scaffold00015	472250	472750	27	['G']	['-']
220	scaffold00015	472250	472750	28	['T']	['-']
221	scaffold00015	472250	472750	29	['T']	['-']
222	scaffold00015	472250	472750	30	['C']	['-']
223	scaffold00015	472250	472750	31	['T']	['-']
224	scaffold00015	472250	472750	32	['G']	['-']
225	scaffold00015	472250	472750	33	['T']	['-']
226	scaffold00015	472250	472750	34	['T']	['-']
227	scaffold00015	472250	472750	35	['G']	['-']
228	scaffold00015	472250	472750	534	['T']	['-', 'C']
229	scaffold00015	472500	473000	232	['T']	['-', 'C']
230	scaffold00015	476750	477250	525	['C']	['-', 'G']
231	scaffold00015	477000	477500	247	['C']	['-', 'G']
232	scaffold00015	481500	482000	497	['G']	['-', 'A']
233	scaffold00015	481750	482250	240	['G']	['-', 'A']
234	scaffold00015	487500	488000	279	['T']	['C', '-']
235	scaffold00015	487750	488250	25	['T']	['C', '-']
236	scaffold00015	489500	490000	411	['T']	['G', 'A', '-']
237	scaffold00015	489750	490250	135	['T']	['A', '-', 'G']
238	scaffold00015	489750	490250	422	['T']	['A', '-']
239	scaffold00015	490000	490500	165	['T']	['A', '-']
240	scaffold00015	499750	500250	482	['G']	['-', 'A']
241	scaffold00015	510000	510500	249	['T']	['-', 'G']
242	scaffold00015	511250	511750	253	['A']	['-']
243	scaffold00015	511250	511750	254	['A']	['-']
244	scaffold00015	511250	511750	255	['G']	['-']
245	scaffold00015	511250	511750	256	['T']	['-']
246	scaffold00015	518250	518750	165	['G']	['-', 'A']
247	scaffold00015	519500	520000	302	['C']	['G', '-', 'A']
248	scaffold00015	521500	522000	434	['A']	['-', 'T']
249	scaffold00015	521500	522000	455	['A']	['-', 'T']
250	scaffold00015	521500	522000	513	['A']	['-', 'C']
251	scaffold00015	521750	522250	134	['G']	['-', 'A']
252	scaffold00015	521750	522250	136	['A']	['-', 'T']
253	scaffold00015	521750	522250	142	['A']	['-', 'T']
254	scaffold00015	521750	522250	199	['A']	['-', 'C']
255	scaffold00015	521750	522250	353	['G']	['-', 'A']
256	scaffold00015	521750	522250	358	['C']	['-', 'T', 'G', 'A']
257	scaffold00015	522000	522500	99	['G']	['A', '-']
258	scaffold00015	522000	522500	104	['C']	['A', '-', 'G', 'T']
259	scaffold00015	526000	526500	474	['G']	['A', '-']
260	scaffold00015	526250	526750	213	['G']	['A', '-']
261	scaffold00015	526750	527250	294	['G']	['A', '-']
262	scaffold00015	527000	527500	6	['G']	['-', 'A']
263	scaffold00015	548500	549000	486	['T']	['C', '-']
264	scaffold00015	556250	556750	401	['T']	['G', '-']
265	scaffold00015	556500	557000	130	['T']	['G', '-']
266	scaffold00015	559750	560250	454	['T']	['G', '-']
267	scaffold00015	560000	560500	196	['T']	['G', '-']
268	scaffold00015	560000	560500	257	['C']	['A', '-', 'T']
269	scaffold00015	560250	560750	368	['T']	['C', '-']
270	scaffold00015	565250	565750	259	['T']	['C', '-']
271	scaffold00015	569250	569750	341	['C']	['-', 'T']
272	scaffold00015	569500	570000	87	['C']	['-', 'T']
273	scaffold00015	570000	570500	261	['A']	['-', 'G']
274	scaffold00015	572500	573000	417	['G']	['-', 'C']
275	scaffold00015	572750	573250	163	['G']	['C', '-']
276	scaffold00015	585750	586250	298	['C']	['G', '-', 'T']
277	scaffold00015	586000	586500	39	['C']	['-', 'T']
278	scaffold00015	589250	589750	282	['G']	['-', 'T', 'C']
279	scaffold00015	592500	593000	429	['C']	['-', 'G', 'T']
280	scaffold00015	594250	594750	405	['G']	['-', 'A']
281	scaffold00015	594500	595000	128	['G']	['-', 'A']
282	scaffold00015	600750	601250	496	['G']	['-', 'C']
283	scaffold00015	601000	601500	239	['G']	['-', 'C']
284	scaffold00015	615500	616000	211	['G']	['-', 'A']
285	scaffold00015	631000	631500	406	['G']	['A', '-']
286	scaffold00015	720750	721250	41	['C']	['-', 'G']
287	scaffold00017	216750	217250	239	['T']	['-', 'G']

288	scaffold00017	217500	218000	296	['T']	['C', 'A', '-']
289	scaffold00017	217500	218000	396	['T']	['A', 'G', '-']
290	scaffold00017	218250	218750	43	['G']	['-', 'A', 'C']
291	scaffold00017	218250	218750	276	['A']	['-', 'C']
292	scaffold00017	218250	218750	279	['A']	['-', 'G']
293	scaffold00017	225500	226000	285	['G']	['-', 'C', 'T']
294	scaffold00017	225750	226250	329	['G']	['A', '-', 'T']
295	scaffold00017	226000	226500	59	['G']	['T', '-']
296	scaffold00017	226000	226500	60	['A']	['G', '-']
297	scaffold00017	226000	226500	380	['C']	['T', '-', 'G']
298	scaffold00017	226250	226750	118	['C']	['-', 'G', 'T']
299	scaffold00017	226250	226750	299	['G']	['-', 'A']
300	scaffold00017	226250	226750	309	['T']	['-', 'C']
301	scaffold00017	226500	227000	41	['G']	['-', 'A']
302	scaffold00017	226500	227000	50	['T']	['-', 'C']
303	scaffold00017	226750	227250	200	['A']	['-', 'T', 'C']
304	scaffold00017	359750	360250	482	['G']	['C', '-']
305	scaffold00017	432500	433000	53	['A']	['G', '-']
306	scaffold00017	432500	433000	252	['T']	['C', '-']
307	scaffold00017	432500	433000	286	['A']	['C', '-']
308	scaffold00017	432500	433000	341	['A']	['G', '-']
309	scaffold00017	432500	433000	342	['A']	['C', '-']
310	scaffold00017	432750	433250	31	['A']	['-', 'C']
311	scaffold00017	432750	433250	86	['A']	['-', 'G']
312	scaffold00017	432750	433250	87	['A']	['-', 'C']
313	scaffold00017	432750	433250	318	['C']	['-', 'G']
314	scaffold00017	433000	433500	64	['C']	['-', 'G']
315	scaffold00017	433250	433750	337	['T']	['G', '-']
316	scaffold00017	433250	433750	368	['A']	['C', '-']
317	scaffold00017	434000	434500	111	['C']	['T', '-']
318	scaffold00017	595250	595750	371	['A']	['G', '-']
319	scaffold00017	595500	596000	251	['A']	['-', 'G']
320	scaffold00017	595500	596000	322	['G']	['-', 'A']
321	scaffold00017	595750	596250	68	['G']	['-', 'A']
322	scaffold00044	33250	33750	245	['C']	['-', 'T']
323	scaffold00044	45000	45500	123	['A']	['-', 'C']

Identified genetic variants specific to the *Triticum* lineage are located in BR0032 genome, giving the scaffold ("BR0032 scaffold"), the genomic fragment coordinates within this scaffold ("fragment start" and "fragment stop") and the position within this 500-base fragment ("position")
"Triticum allele": nucleotidic base found in all *Triticum* lineage isolate genomes in this position ;
"non-*Triticum* allele": base(s) found in other genomes in this position

Table S4. List of primers and probes used in this study.

Group number	Primer type	Primer name	BR0032 scaffold	BR0032 fragment	Sequence (5'-3')
C1	reverse	R1	15	373000	ACTCGATACAGATATTCGGTTCGTC
C2	reverse	R2	15	373000	CCGTGCTAATTAGGCCAGGC
C3	forward	F3	15	373750	ACTTCAAAGGTCAGGATCCAACA
C3	reverse	R3	15	373750	TACTGGGGCAACGAACCG
C4	forward	F4	15	374000	TTCGTTGCCCCAGTACAGC
C4	reverse	R4	15	374000	CGAACTTAGCTACTAAACCAACGG
C5	**reverse**	**R5**	**15**	**402500**	**TGCGCATGGCTTGCTTG**
C6	forward	F6	15	428000	TGTGCAACTACTACGTATGTCCG
C6	reverse	R6	15	428000	GTTTTTATGCTGCTGTATTTGGGTTC
C8	forward	F8	15	373250	ACGATAAGTGAGGAGGTGCATAC
C9	**forward**	**F9**	**15**	**373250**	**ACCCAATTAGTAAAAGTGCCTTGTC**
C10	forward	F10	15	373250	GCATACCCAATTAGTAAAAGTGCCT
C11	forward	F11	15	394250	TGTTTTGATCAATGGCCGGATTT
C11	reverse	R11	15	394250	CGGGCTTACCATAGTAAAGTCCTAT
C12	reverse	R12	15	402500	CATGGCTTGCTTGCCTTTTTTT
C13	reverse	R13	15	402500	CTGCGCATGGCTTGCTTG
C14	forward	F14	15	402750	CCCTGCGTGAAAAAAAGGCAA
C14	reverse	R14	15	402750	GTTGTTGGTGAAAGAAAGCTACTCG
C15	forward	F15	15	428000	AGCTTCCTCTTCAGTGCGAC
C15	reverse	R15	15	428000	TCATTGAGCGGACGGACATAC
C16	forward	F16	15	511250	CAAGATGCACCATTCTAAACTGGA
C16	reverse	R16	15	511250	GTTAGCGATTATCTATTTCACCACTTACTT
C17	**forward**	**F17**	**15**	**511250**	**CGATAGAAACTTGAGGAAGATCAAGTAAG**
C17	**reverse**	**R17**	**15**	**511250**	**TCACCGAGAGATGTGCCAC**
C17	**probe**	**P17**	**15**	**511250**	**TCGCTAACAATGTCCACCCCGCC**
C18	forward	F18	15	521750	CCTCTGCATTTTTACCCATCGAG
C18	reverse	R18	15	521750	GCAGCGGGGAACATGGAT
C19	**forward**	**F19**	**17**	**432500**	**CTGCTCAATGGCCCCAGAT**
C19	**reverse**	**R19**	**17**	**432500**	**CTCGCCCAGCTGAAGCTT**
C21	forward	F21	17	432500	CCAGCCCGCAGATGTCAAAA
C21	reverse	R21	17	432500	CTGGGCGTGGAATTCGTCA
C22	forward	F22	17	226000	CAACAAAGCAAGTCGCAAGGA
C22	reverse	R22	17	226000	GAATAGTTCTTGAGGGAACCAACAC
C23	forward	F23	17	218250	CCGAGAACCCCGCACTA
C23	reverse	R23	17	218250	ACGGCGACCATGGACTT
C24	forward	F24	17	595500	AAGGGTTTGATAGAAGATGGGATGA
C24	reverse	R24	17	595500	GAGGGAGATGGACTGCTGAC
C25	forward	F25	17	595500	AGGGCATCATTTTGATGGGCATA
C25	reverse	R25	17	595500	TCATACTAAAAGGATAAGTTGGCGAATTC
C26	forward	F26	44	45000	CAGGTGTCCCAATATCAAGCACTA
C26	reverse	R26	44	45000	GAATCTATGGCTAACACGGAAACTTA
C28	forward	F28	10	1275750	AGGATGGCGCACCACTG
C28	reverse	R28	10	1275750	CGTTTCCGGTCTCGACGA
C29	forward	F29	15	521750	TGTTTCCTTTGTGCCTCTGCA
C29	reverse	R29	15	521750	TATCATCCCCGCAAACAGCC
C30	forward	F30	15	521750	GAGTAGTGTTGGAATCGAAGTTCAC
C30	reverse	R30	15	521750	TGTTTGAGCGCGATCCGT
C31	forward	F31	15	560000	TGAGTCAACATCAGGCCGAAT
C31	reverse	R31	15	560000	CAGCAAAGAGAACGGTTCCCA
C32	forward	F32	15	560000	CTGTATGTGCTCTATGTCTTTAGTCTC
C32	reverse	R32	15	560000	AGTGGAAGTGGGAATGGGGTA
C33	forward	F33	15	631000	ATTCTTGGCCCCACTTTTTTCTG
C33	reverse	R33	15	631000	GCATTTTGTGCTGCTAAACGAAC
C34	forward	F34	15	350750	TCTGTCTATATGTCACACGAAAATGAG
C34	reverse	R34	15	350750	GGATCAGCGAGTTTGCACTG
C35	forward	F35	15	350750	TCTTGACAGCTCCAGGCAAAT
C35	reverse	R35	15	350750	TAGATCGCGTCATGTTGGACC
C1 & C2	forward	F1 & F2	15	373000	CATTCTCGGTTTTGCATTAAGTCTAGAC
C19 & C20	**probe**	**P19 & P20**	**17**	**432500**	**GCGGCCGACAAACCTGCTGC**
C5 & C12 & C13	**forward**	**F5 & F12 & F13**	**15**	**402500**	**AACCACGGCCATTGAGGTAAT**
C5 & C12 & C13	**probe**	**P5 & P12 & P13**	**15**	**402500**	**ACAGGGATGTATTGTCGCGATGT**
C8 & C9 & C10	**probe**	**P8 & P9 & P10**	**15**	**373250**	**ACTGTGTGGATTCAAAAAGTTAACCTTGACC**
C8 & C9 & C10	**reverse**	**R8 & R9 & R10**	**15**	**373250**	**CCCATAAGCTTATTCAGTGCGG**

Article 3: A PCR, qPCR and LAMP toolkit for detection of the wheat blast pathogen in seeds

Les principaux objectifs de cet article étaient de :

- Elargir la recherche de polymorphismes spécifiques de la lignée Triticum en explorant de nouvelles régions génomiques afin de d'éliminer les amplifications non-spécifiques obtenues avec le test C17 ;
- Proposer des tests de diagnostic utilisant de multiples techniques pour s'adapter à différentes contraintes financières, matérielles, de temps ou de main-d'œuvre ;
- Développer un protocole permettant l'utilisation de ces tests pour la détection de l'agent pathogène sur grains, et évaluer leur sensibilité.

Ce travail a été fait en collaboration avec Axel Chatet, stagiaire en master 2. Dans ce projet, j'ai pris en charge la partie bio-informatique, j'ai développé les amorces utilisées pour l'amplification isothermale LAMP et j'ai participé à l'encadrement du stage.

A PCR, qPCR and LAMP toolkit for detection of the wheat blast pathogen in seeds

Thierry, Maud*; Chatet, Axel*; Fournier, Elisabeth; Tharreau, Didier; Ioos, Renaud
* both authors equally contributed to this work.

INTRODUCTION

Wheat blast is an emergent disease threatening the global wheat production and therefore the global food safety. The major agricultural impact of the disease and its fast propagation require an accurate and fast identification protocol of the pathogen in order to prevent its spread and avoid further exchange of contaminated biological material at global scale.

The pathogen responsible for wheat blast (*Pyricularia oryzae* Triticum lineage, synonymous *Magnaporthe oryzae)* is capable of infecting all aerial parts of the wheat plant, but spike infection is the most common symptom observed in the field (Islam et al., 2019; Pieck et al., 2017). Symptoms of wheat blast include necrotic lesions on leaves, stems and grains, partial or total bleaching of the spikes leading to sterility or empty grains. Wheat blast disease reduces grain yield and grain quality (Islam et al., 2019; Urashima et al., 2004). During major epidemics, it caused up to 100% of yield loss (Cruz and Valent, 2017). This disease has a strong potential economic impact since cereals account for about 40% of the world's agricultural yield, wheat coming first to rice and maize, with about 700 million annual tons consumed by humans (FAOSTAT: http://www.fao.org/faostat/fr/#home).

This devastating disease first emerged in 1985 in the state of Paraná in Brazil (Igarashi et al., 1986). Then, the disease spread in the neighbouring states of São Paulo, Mato Grosso do Sul in 1986 and Rio Grande do Sul in 1987 causing a decrease in the yield of 95% of wheat crops in the Cerrado (Anjos et al., 1996; Goulart, 1992; Goulart et al., 2007; Picinini and Fernandes, 1990). The pathogen subsequently spread to eastern Bolivia in 1996, eastern Paraguay in 2002 and northern Argentina in 2007, before being transported to Bangladesh in 2016 (Cruz and Valent, 2017).

P. oryzae spores are not able to spread far with the wind (Urashima et al., 2007). However, the transport of contaminated seeds facilitates the spread of the fungus over a long distance (Gomes et al., 2018). A study demonstrated, by a comparative genomic analyse, a close relationship between isolates collected from wheat in different parts of Bangladesh and of *P. oryzae* found in the epidemic zone in South America (Islam et al., 2016; Malaker et al., 2016). The transport of grains contaminated with *P. oryzae*, harvested in the wheat blast epidemic zone in Brazil, likely caused the recent emergence of this pathogen in Bangladesh in 2016 (Islam et al., 2016).

P. oryzae is responsible for blast disease on numerous Poaceae species (Klaubauf et al., 2014; Ou, 1985). Phylogenetic analyses of 81 genomes of *P. oryzae* isolates sampled from 12 different genera of Poacea revealed divergent multiple lineages within *P. oryzae*, each of which was mostly associated with one specific host plant genus (Gladieux et al., 2018). The existence of these host-specific lineages revealed an incipient speciation following host jump or host range expansion of the pathogen. However, the genetic divergence (number of differences per kilobase) observed between host-specific lineages was less than 1% on the entire genome - in comparison, the genetic diversity between different species of Pyricularia (*P. grisea, P. oryzae, P. pennisetigena*) is greater than 10% - and gene flow was detected between host specific lineages (Gladieux et al., 2018). *P. oryzae* therefore represents a single species grouping different host-specific lineages.

The majority of the isolates sampled on infected wheat are clustered within one of these host-specific lineages, the Triticum lineage. However, *P. oryzae* isolates are sometimes able of opportunistic infection on a host plant different from their original host, but causing much lower symptoms. Several studies recorded opportunistic infection on wheat caused by isolates belonging to the Lolium lineage of *P. oryzae*, which is the genetically closest host-specific lineage to the Triticum lineage. These isolates, however, were weakly aggressive on wheat during artificial inoculation tests and do not appear to be capable of causing major epidemics on this host (Farman et al., 2016; Pieck et al., 2017). The Triticum lineage of *P. oryzae* is therefore held responsible for wheat blast epidemics.

Accurate and fast methods for detecting wheat blast isolates are required to limit or prevent the spread of the pathogen in disease-free areas (Cruz and Valent, 2017). A misidentification of the pathogen could lead to drastic measures such as the unjustified destruction of seeds or healthy biological material. On the other hand, a false-negative result could result in a wheat blast outbreak in a new geographical area.

However, intra-specific detection is challenging, since gene flow likely regularly occurs among lineages or subpopulations belonging to the same species. The detection method must be capable of discriminating isolates responsible for wheat blast epidemics (belonging to the Triticum lineage) of isolates belonging to the other host-specific lineages of the species but which may be capable of opportunistic infections on wheat plants. Host-specific lineages have an identical morphology in pure culture which does not allow to differentiate them visually (Thierry et al., 2019). Pathotyping tests can be used for diagnostic but are time-consuming and cannot differentiate unambiguously

opportunistic infections. DNA-based detection tests are a good alternative because they allow detection at very precise taxonomic levels. Provided, however, to identify a specific polymorphism of the targeted taxon. In the case of wheat blast, the identification of such polymorphisms is made difficult by the low genetic divergence and the gene flow between host-specific lineages (Gladieux et al., 2018).

Currently, several DNA-based diagnostic tests exist for the detection of wheat blast isolates (Pieck et al., 2017; Thierry et al., 2019; Yasuhara-Bell et al., 2018). All these tests are highly, but not perfectly, inclusive (i.e. detecting all wheat-blast isolates) and specific (not detecting non-wheat-blast isolates): none of them allow an optimal detection of the pathogen. Pieck et al., (2017) and Yasuhara-Bell et al., (2018) developed PCR, qPCR and LAMP diagnostic tests targeting the same genomic region, the MoT3 locus. This region was selected because highly conserved in wheat-blast isolates and absent in most non-wheat-blast isolates. However, some wheat blast isolates, such as the BR0032 isolate, do not include the MoT3 locus and remained undetectable using these tests. Thierry et al., (2019) developed the C17 qPCR test targeting a different genomic region. This test allowed the detection of all wheat blast isolates tested so far. However, some isolates non-pathogenic on wheat were also positively detected with this test leading to 4% of false positive results.

The objectives of this work were to (i) design primers targeting new genomic regions in order to identify polymorphisms fully specific to the Triticum lineage; (ii) develop a toolkit of detection tests using multiple DNA amplification techniques (PCR, qPCR and LAMP) to be suitable for any type of analysis and improving current wheat blast detection; (iii) verify the capacity of these tests to detect the pathogen on artificially contaminated wheat grains.

RESULTS

Primers screening using a small DNA panel: search for specific polymorphisms

Inclusivity (i.e. positive result for all targeted isolates) and specificity (no detection of non-targeted isolates) are the most important features for a detection test. All designed primer pairs were challenged with DNAs extracted from wheat-borne *P. oryzae* isolates (isolates sampled on wheat) and non-wheat-borne *P. oryzae* isolates (isolates sampled on other poaceae than wheat) to assess their inclusivity and specificity.

PCR primers. Forty-nine designed primer pairs were first screened using conventional PCR on a small DNA panel (composed of five wheat-borne isolates and ten non-wheat-borne isolates).

Among them, 12 primers pairs displayed a full inclusivity and specificity on this small panel (Supplementary figure 1). Four of these last (C45, C74, C82 and C92) were selected for further analyses. The four primers pairs are designed on three different scaffolds of the reference genome BR0032 (scaffolds 15, 17 and 44).

LAMP. Five groups of primers were designed for a LAMP isothermal amplification targeting 3 loci. The inclusivity and specificity of the five groups of LAMP primers designed were tested on a small panel composed of three wheat-borne isolates and four non-wheat-borne isolates (Supplementary figure 2). None of the five groups of primers tested displayed a full specify and inclusivity. Two groups of primers did not amplify any of the DNAs (group 2 and group 3) and two groups amplified every DNA tested (group 1 and group 4). Group 5 was selected for further analyses because primers of this group allowed to amplify every wheat-borne isolates in a very short time (around three minutes), even if full specificity of this primer group could not be achieved.

Primers specificity and inclusivity assessment using a large DNA panel

Inclusivity and specificity of the primers previously selected were assessed using a larger DNA panel of 113 or 185 strains (Figure 1).

PCR and qPCR primers: Inclusivity and specificity of C45, C74, C82 and C92 primer pairs were first assessed by conventional PCR (Figure 1). The primers C45 displayed the highest inclusivity and specificity and was further tested on additional DNAs by conventional PCR and by real time PCR by adding a fluorescent probe.

The C45 primers are designed in the scaffold 15 of the reference genome BR0032 and target a single SNP positioned at the 3' end of the forward primer. PCR and qPCR tests using these primers and probe displayed 97% inclusivity. All wheat-borne isolates but one, isolate AG0102, were amplified. To validate AG0102 host spectrum, a pathogenicity test was done by inoculating a solution of AG0102 spores on the leaves of the susceptible wheat cultivar Thésée. No symptom was observed, questioning the virulence of this isolate on wheat. Furthermore, no amplification of AG0102 DNA was observed using either the MoT3 test (Pieck et al., 2017) or the C17 test (Thierry et al., 2019).

C45 primers displayed a full specificity by conventional PCR and no non-wheat-borne isolates were amplified. However, late amplifications of non-wheat-borne isolates appeared when using real time PCR. Additional repetitions of the qPCR test on these DNAs evidenced that these amplifications were not fully repeatable and only occurred in some of the repetitions. Finally, qPCR test using C45

primers allowed amplification of all wheat-borne isolates except AG0102 (Ct values between 23 and 28), but 10 out of 151 non-wheat-borne isolates were sometimes amplified with late Ct values (between 35 and 39).

LAMP primers: LAMP primers group 5 succeeded to amplify all wheat-borne isolates DNA, except the AG0102 isolate, in a very short time (between 3:37 and 4:10 minutes). However as expected a full specificity could not be achieved. Six isolates (AG0067, Cd88215, CR0023, CR0057, JP0031, JP0033) out of 151 (4%) were amplified with time values equivalent to the values observed with wheat-borne isolates, making them indistinguishable. Three of these six were also positively detected using the C17 qPCR test (Thierry et al., 2019). Late and often non repeatable amplifications could also be observed after 9 to 25 minutes of amplification (Figure 1).

Primers sensitivity

C45 qPCR and PCR primers: Sensitivity for PCR and qPCR primers were assessed using serial 10-fold dilutions of genomic DNA of three *P. oryzae* wheat-borne isolates (BR0031, BL0063, BL0023) and 10-fold dilutions of plasmidic constructions integrating the sequences targeted by the primers. PCR and qPCR primers C45 succeeded to detect up to 5 pg of DNA/reaction and up to 12 plasmidic copies/reaction for genomic DNA and plasmidic DNA, respectively. This limit of detection was validated using 16 replicates for each DNA concentration. PCR reaction efficiency was measured using the plasmidic dilution to 92% and the high R^2 value (R^2=0.9969) evidenced a good correlation between initial plasmidic DNA quantity and the Ct values.

Group 5 LAMP primers: LAMP primers group 5 sensitivity was assessed by serial 10-fold dilutions of genomic dilution of two *P. oryzae* wheat-borne isolates (BR0031 and BL0023). LAMP primers succeeded to amplify up to 5 pg of genomic DNA per reaction. This limit of detection was validated using 5 replicates.

Pathogen detection in contaminated seeds

We assessed the ability of the PCR (C45 primers), qPCR (C45 primers + probe) and LAMP (group5 primers) tests to detect the pathogen in a seed matrix. Wheat seeds were artificially contaminated with a spore suspension of the pathogen. Several grain lots of 400 seeds in total were constituted containing 1, 2, 5, 10, 15, 20 or 50 artificially contaminated seeds mixed with non-contaminated seeds. Total DNA were extracted from each lot after blending. Before the blending step, some lots were incubated 72h in potato dextrose broth media to test if this enrichment process improves pathogen detection.

Without the incubation step, the conventional PCR succeeded to amplify all nine replicates in seed lots containing 15 contaminated seeds or more. The real-time PCR test was less sensitive because only succeeded to amplify all nine replicates in seed lots containing 50 contaminated seeds or more. LAMP tests on the other hand did not succeed to amplify all replicates for any of the seed lots tested.

The incubation step highly improved the detection for all tests. After 72h of incubation, the PCR test, as well as the qPCR test, were able to detect the pathogen, for all replicates, in seed lot containing only one contaminated seed. The Ct values of the qPCR test were around 34. The LAMP primers allowed the detection of the pathogen after an enrichment step, for all replicates, in lot containing 2 or more contaminated seeds in less than 5 minutes.

DISCUSSION

Wheat blast has a major agricultural and economic impact. Reliable detection tests are required to avoid the propagation of this emergent disease. Intraspecific detection of the Triticum lineage of *P. oryzae* responsible for wheat blast is, however, difficult to achieve. The strong genetic proximity and the gene flow between the host-specific lineages of *P. oryzae* makes it difficult to detect polymorphisms fully specific of the Triticum lineage and no detection test developed to date displayed perfect specificity and inclusivity (Pieck et al., 2017; Thierry et al., 2019; Yasuhara-Bell et al., 2018).

In this study we designed primers targeting new genomic regions in order to find polymorphisms with greater specificity for the *P. oryzae* Triticum lineage. The screening of these primers highlighted the C45 pair. The specificity of this pair is based on a single substitution located at the 720791 position of the scaffold 15 in the BR0032 reference genome. The isolates belonging to the Triticum lineage display an “C” at this position while this nucleotide is substituted by a “G” in other lineages. This single substitution allowed the C45 pair to amplify the DNA of all the isolates sampled on wheat, with the exception of isolate AG0102 whose pathogenicity on wheat is questioned (see result section), while no amplification of isolate sampled on other poaceae using conventional PCR. Such a specificity was not obtained to date.

The use of these primers supplemented with a fluorescent probe allowed the development of a real-time PCR test. Real-time PCR has two major advantages compared to conventional PCR, (i) it quantifies the presence of the pathogen in the samples tested and (ii) is highly sensitive, allowing

generally the detection of a lower amount of pathogen DNA. The C45 qPCR test allowed the same level of inclusivity as conventional PCR (all isolates from wheat were amplified with the exception of AG0102). However, we have been confronted with late non-specific amplifications. Ten isolates showed amplifications, for some of the repeats, at Ct greater than 35.50. The specificity of the targeted polymorphism is not questioned by these results. The strong shift of Ct values between amplification of isolates sampled on wheat and the non-specific amplifications clearly demonstrates a difference in their nucleotide sequence. However, during the amplification cycle a partial degradation of the primers could suppress the single nucleotide carrying the specificity and lead to these late amplifications. Late amplification obtained using the C45 qPCR test should be validated using other previously developed tests (MoT3 (Pieck et al., 2017), C17 (Thierry et al., 2019)) or by sequencing the locus. Indeed, the combination of several tests allows the validation of the results because different genomic regions are targeted by these tests leading to different pattern of amplification (different false positives or false negative results).

The single substitution targeted by C45 did not allow the development of LAMP primers. So, the LAMP primers (group 5) were targeted to another genomic region located at the 428000 position of the BR0032 scaffold 15. This test detected all isolates sampled on wheat, except isolate AG0102, in a very short time (<4 minutes). However, specificity assessed on 185 DNAs of different *P. oryzae* isolates showed several nonspecific amplifications. Six isolates (AG0067, Cd88215, CR0023, CR0057, JP0031 and JP0033) led to amplifications at similar times as the target isolates (Triticum isolates), between 3 and 4 minutes. Twenty-two other nonspecific amplifications were observed later, after 8 minutes of amplification. In order to eliminate these last non-specific amplifications, the LAMP PCR run can be stopped after 8 minutes without impacting the sensitivity of the test. However, the six non-specific amplifications observed between 3 and 4 minutes will not be eliminated by these modifications but positive results can be confirmed with the C45 PCR or qPCR test which does not amplify these non-target isolates DNA. The LAMP group 5 test could therefore guarantee a result in 8 minutes. If the result is positive on all replicates, then confirmation by PCR or qPCR should be done.

Knowing the sensitivity of a test is essential to validate the method. The limit of detection was very low for the three tests. C45 PCR and qPCR tests as well as LAMP group 5 tests were capable of detecting as little as 5 pg of DNA per PCR reaction. The sensitivity of the PCR and qPCR C45 tests was also evaluated using plasmid DNA. The limit of detection (T + LOD) obtained for both tests was 12 plasmid copies of the target DNA per PCR tube.

Detection of the pathogen on wheat grains is essential to validate the health status of seeds before movement at importation sites. However, amplification inhibitors may be present in seeds and impact the performance of detection tests (Mancini et al., 2016). An enrichment phase in fungal biomass is often used to overcome these barriers (Mancini et al., 2016). In our study, only the PCR and qPCR C45 tests allowed repeatable amplification without enrichment phase. Surprisingly, the C45 PCR test showed better sensitivity when used in seed matrix (detection of 15 contaminated seeds per 400-seed lot for all replicates) than the qPCR test (detection of 50 contaminated seeds per 400-seed lot). The addition of an enrichment phase consisting of the incubation of grains for 72 hours in a rich medium greatly improved the sensitivity of all the tests. Both the PCR and qPCR C45 tests allowed the systematic detection of the pathogen in a batch of 400 seeds containing only one artificially contaminated seed which represents a 0.25% infection rate. The LAMP group 5 test allowed the systematic detection of the pathogen in lots containing 2 contaminated seeds after the enrichment phase. Additionally, none non-specific late amplifications were obtained on seed lots highly contaminated by the BR0079 isolate (isolate sampled on Eleusine indica). Seed lots containing 20 and 50 seeds contaminated with BR0079 spores do not cause any PCR or qPCR signal even after an incubation period of 72h.

In order to use the detection tests developed in this study to test the presence of *P. oryzae* isolates belonging to the Triticum lineage on wheat seeds, lots of 400 seeds must be formed and incubated for 72 hours in a culture medium (PDB). DNA extracted from these grains will then serve as a template for detection tests. The detection method used may be selected depending on the financial, material, time and labor resources. However, we recommend in the first place to use fully inclusive detection tests, which do not generate false negatives on the isolates tested to date (C17, C45 PCR, C45 qPCR, LAMP group5 tests). The conventional C45 PCR test, however, is the only one that displayed a perfect specificity for the isolates tested. Given the high genetic proximity of the host-specific lineages as well as the potential gene flow between these lineages, all positive results must be validated by the use of another test to avoid costly and unjustified quarantine measures such as what happened for Fusarium oxysporum f. sp. cubense race 4 (Magdama et al., 2019).

MATERIAL ET METHODS

Biological material

P. oryzae DNA

Wheat blast epidemics are most likely caused by isolates belonging to the Triticum lineage of the *Pyricularia oryzae* species (Gladieux et al., 2018; Thierry et al., 2019). The goal of this study is to specifically detect these isolates. 185 *P. oryzae* isolates were analysed to assess the specificity and

the inclusivity of the diagnostic tests designed in this study. Among these *P. oryzae* isolates, 34 were sampled on infected wheat (Triticum sp.) and 151 were sampled on 28 different poaceae genus. Isolates sampled on wheat are referred as wheat-borne isolates in this manuscript. Isolates were sampled in various countries in order to maximize the genetic diversity of the pathogen. 19 isolates belonging to 9 other fungal species pathogenic on wheat were also included (Supplementary table 1). All DNA concentration were measured with a spectrophotomer (nanodrop 200- thermo Fisher Scientific) and adjusted at 0.5ng/µl by dilution with a Tris EDTA 1X buffer. Amplifiability of all DNAs was validated by the FungiQuant real time PCR assay targeting a conserved 351 bp region in the fungal 18S rRNA gene (Liu et al., 2012).

Positive control plasmids

Positive controls for PCR and qPCR C45 detection tests were prepared by cloning the targeted sequence. Plasmids are deemed stable, homogeneous during pipetting, easily quantifiable and producible in virtually unlimited quantities. The genomic region targeted by the C45 primer pair is amplified by PCR and the size of the amplicon is checked on electrophoresis gel. The sequence is inserted into the pCR4-TOPO vector following the TOPO TA cloning kit protocol (Invitrogen) and used to transform TOPO10 chemically competent bacteria (Escherichia coli) according to the manufacturer's instructions. After culturing at 37 ° C and selecting by qPCR the clones with a fragment integration, the plasmids are purified using the Nucleospin® (Macherey-Nagel) plasmid kit. The plasmid solution is used as a positive control of the PCR and qPCR reactions of the c45 couple.

Seeds inoculation and DNA extraction

The wheat-borne isolate BL0092 was cultured on potato dextrose agar (PDA) medium at 23 ° C, under day-night alternance (12H/12H). After 7 days of culture, fungal spores were collected by adding 5ml of sterile water to the petri dish and scratching the mycelium to take off the spores. The resulting solution is filtered to remove mycelium fragments and conserve only the spores. Finally, the concentration of the solution was calibrated using a haemocytometer at 200 spores/µL, supplemented by two drops of tween 20 and used to artificially inoculate wheat seeds of the susceptible wheat variety Filon. These seeds were previously disinfected by a stay of 10 minutes in a 1.5L solution of sodium hypochlorite diluted at 1.5°. The inoculation was carried out by depositing 10µL of spore solution on each grain and drying the seeds overnight in a sterile atmosphere. Multiple lots of 400 seeds in total but containing different proportions of contaminated seeds were created (0, 1, 2, 5, 10, 15, 20 or 50 artificially contaminated seeds per lot). Fungal enrichment was realized by incubating the lots in 15ml of potato dextrose broth (PDB) media during 72h at room

temperature (23°C). The grinding of incubated or non-incubated seed lots was carried out with 30 mL of PDB using the Microtron ™ MB 550 Laboratory Mixer (Kinematica ™) until a milky mixture was obtained. For each sample, three samples of 500µL each were taken using a truncated 1mL cone and used for DNA extraction with the Nucleospin® Plant II (Macherey-Nagel) extraction kit following the manufacturer's instructions. To serve as a negative control, the same spore inoculation protocol was performed with the BR0079 isolate sampled on Eleusine indica and the DNA of two seed lots containing respectively 20 and 50 inoculated seeds was extracted after a fungal enrichment step.

Primers design

PCR and qPCR: The primers and probes are designed to target Triticum-specific polymorphisms identified by comparison of 76 *P. oryzae* genomes including 20 genomes assigned to the Triticum lineage. Here we have studied some of the polymorphisms highlighted but not exploited in the publication of Thierry et al., (2019). In addition, less stringent parameters applied in the bioinformatic pipeline allowed us to identify new polymorphisms specific to the Triticum lineage. Primers and probes for PCR and qPCR were designed using the Geneious 11.1.2 tool. The specific polymorphisms of the Triticum lineage were positioned in the 3' end of the forward and reverse primers to maximize the specificity of the primers. The forward primer of the C45 pair was additionally purified by HPLC by the manufacturer to ensure the integrity of the primer, whose specificity is based on a single nucleotide located in the 3’ end of the forward primer.

LAMP: The four primers required for each LAMP reaction including the two external primers F3 and B3 and the two internal primers FIP (F1c + TTTT + F2) and BIP (B1c + TTTT + B2) were designated with the PrimerExplorer V5 online tool (https://primerexplorer.jp/e/).
All the primers used in this study were synthesized by Eurogentec. The list of all the candidate primers is available in Supplementary table 2.

Primers screening

PCR primers: All primers pairs designed were screened using a small DNA panel from 15 isolates including five isolates sampled on Triticum (BR0086, BR0036, AG0103, BL0017 and BL0093), three on Oryza (CH1120, FR0013 and BR0019), four on Lolium (AG0064, CHW, PL2-1 and CR0057), one on Eleusine (IN0113), one on Echninochloa (CR0023) and one on Eriochloa (JP0033). The PCR reactions for the screening were carried out in qPCR condition, without hydrolysis probe, using the LightCycler® 480 Probes Master (Roche) with the following reaction mixture: 1X of the Roche Kit premix qPCR, 0.3µM of primer forward and reverse, 2µL of template DNA (0.5ng / µL) and ultrapure

water for a final volume of 20µL. The PCR reaction is carried out in the LightCycler® 480 II thermal cycler (Roche). The PCR cycle contains (i) initial denaturation step at 95°C for 10 min; (ii) followed by 40 denaturation cycles at 95°C for 15 sec and hybridization-synthesis at 62°C for 55 sec; (iii) followed by a final elongation at 62°C for 10 minutes.

LAMP primers: The small DNA panel used to screen LAMP primers was composed of seven DNA including the DNA of three isolates sampled on Triticum (BR0036, BR0031 and BR0088), one eriochloa (JP0033), one bromus (AG0061), one eleusine (IN0113) and one lolium (CR0057). The mix and amplification cycle used for this screening is same as described in the "PCR, qPCR and LAMP amplifications" section.

PCR, qPCR and LAMP amplifications

Conventional PCR

Conventional PCR mix included 1X polymerase buffer, 5mM MgCl2, 0.25 mM of each dNTP, 0.025 U/µL of HotGoldStar DNA polymerase (HGS Diamond Taq® DNA polymerase), 0.3µM of each forward and reverse primers, 2µL of DNA matrix (fixed at 0.5ng/µL) and ultrapure water to reach the final 20µL volume. The amplification cycle was 1) 10 minutes of initial denaturation step at 95°C, 2) followed by 40 cycles each composed of 30 seconds of denaturation at 95°C, 30 seconds of hybridization at 65°C and 45 seconds of extension at 72°C, 3) a final extension at 65°C during 10 minutes. PCR amplifications were visualized after a one-hour electrophoresis at 110 volts on a 1.5% agarose gel (3g of agarose mixed in 200 mL of TBE 1X buffer). Conventional PCR run were achieved in a LightCycler® 480 II thermal cycler (Roche).

Real time PCR

The real time PCR were achieved using the No ROX mastermix (Eurogentec) using the following mix: 1X of mastermix, 0.3µM of forward and reverse primers, 0.1µM of TaqMan fluorescent probes, 2µL DNA matrix (fixed at 0.5ng/µL) and ultrapure water to reach the final volume of 20µL. The real time PCR was achieved using a RotorGene 4.4.1 (Qiagen). The qPCR amplification cycle is composed of 1) An UNG glycosylase activation step at 50°C during 2min, 2) An initial denaturation step at 95°C during 10min, 3) 40 cycles composed of 15 seconds of denaturation at 95°C and 55 seconds of hybridation-synthesis at 65°C. The detection threshold was manually fixed at 0.02.

LAMP

LAMP amplifications were achieved using the Isothermal Master Mix ISO-001 (OptiGene) with the following mix : 0.2µM of each external primers (F3 et B3), 0.8µM of each internal primers (FIP et

BIP), 1X of isothermal mastermix, 2µL of DNA matrix (fixed at 0.5ng/µL) and ultrapure water to reach the final 25µL volume. LAMP is an isothermal amplification technic. Amplifications were realized using a RotorGene 4.4.1 (Qiagen) thermocycler whose temperature was set at 65°C during 30 minutes. The detection threshold was manually fixed at 0.02.

LITERATURE CITED

Anjos, J., Da Silva, D.B., D'Avila Charchar, M.J., and Rodrigues, G.C. (1996). Occurrence of blast fungus (Pyricularia grisea) on wheat and rye in the savanna region of central Brazil. 31, 79–82.

Cruz, C.D., and Valent, B. (2017). Wheat blast disease: danger on the move. Trop. Plant Pathol. 42, 210–222.

Farman, M., Peterson, G., Chen, L., Starnes, J., Valent, B., Bachi, P., Murdock, L., Hershman, D., Pedley, K., Fernandes, J.M., et al. (2016). The Lolium Pathotype of Magnaporthe oryzae Recovered from a Single Blasted Wheat Plant in the United States. Plant Dis. 101, 684–692.

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J.L.N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., et al. (2018). Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the Rice Blast Fungus Magnaporthe oryzae. MBio 9, e01219-17.

Gomes, D.P., Rocha, V.S., Rocha, J.R. do A.S. de C., Pereira, O.L., and Souza, M.A. de (2018). Potential of transmission of Pyricularia graminis-tritici from plant to seed and from seed to seedling in wheat genotypes with different degrees of blast resistance. J. Seed Sci. 40, 16–24.

Goulart, A. (1992). Incidência da brusone (Pyricularia oryzae) em diferentes cultivares de trigo (Triticum aestivum) em condições de campo. Fitopatol. Bras. 17, 321–325.

Goulart, A.C.P., Sousa, P.G., and Urashima, A.S. (2007). Damages in wheat caused by infection of Pyricularia grisea. Summa Phytopathol. 33, 358–363.

Igarashi, S., Utiamada, Kasuma, and Lopez (1986). Pyricularia sp. em trigo. Ocurrencia de Pyricularia sp no Estado do Parana. Fitopatol. Bras.

Islam, M.T., Croll, D., Gladieux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, Md.S., Gupta, D.R., Rahman, Md.M., Mahboob, M.G., et al. (2016). Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae. BMC Biol. 14, 84.

Islam, M.T., Kim, K.-H., and Choi, and J. (2019). Wheat Blast in Bangladesh: The Current Situation and Future Impacts. Plant Pathol. J. 35, 1–10.

Klaubauf, S., Tharreau, D., Fournier, E., Groenewald, J.Z., Crous, P.W., de Vries, R.P., and Lebrun, M.-H. (2014). Resolving the polyphyletic nature of Pyricularia (Pyriculariaceae). Stud. Mycol. 79, 85–120.

Liu, C.M., Kachur, S., Dwan, M.G., Abraham, A.G., Aziz, M., Hsueh, P.-R., Huang, Y.-T., Busch, J.D., Lamit, L.J., Gehring, C.A., et al. (2012). FungiQuant: A broad-coverage fungal quantitative real-time PCR assay. BMC Microbiol. 12, 255.

Magdama, F., Monserrate-Maggi, L., Serrano, L., Sosa, D., Geiser, D.M., and Jiménez-Gasco, M. del M. (2019). Comparative analysis uncovers the limitations of current molecular detection methods for Fusarium oxysporum f. sp. cubense race 4 strains. PLOS ONE 14, e0222727.

Malaker, P.K., Barma, N.C.D., Tiwari, T.P., Collis, W.J., Duveiller, E., Singh, P.K., Joshi, A.K., Singh, R.P., Braun, H.-J., Peterson, G.L., et al. (2016). First Report of Wheat Blast Caused by Magnaporthe oryzae Pathotype triticum in Bangladesh. Plant Dis. 100, 2330–2330.

Mancini, V., Murolo, S., and Romanazzi, G. (2016). Diagnostic methods for detecting fungal pathogens on vegetable seeds. Plant Pathol. 65, 691–703.

Ou, S.H. (1985). Rice Diseases (IRRI).

Picinini, E.C., and Fernandes, J.M.C. (1990). Occurrence of wheat blast (Pyricularia oryzae) in commercial fields in the state of Rio Grande do Sul, Brazil. Fitopatol. Bras. Braz.

Pieck, M.L., Ruck, A., Farman, M.L., Peterson, G.L., Stack, J.P., Valent, B., and Pedley, K.F. (2017). Genomics-Based Marker Discovery and Diagnostic Assay Development for Wheat Blast. Plant Dis. 101, 103–109.

Thierry, M., Gladieux, P., Fournier, E., Tharreau, D., and Ioos, R. (2019). A genomic approach to develop a new qPCR test enabling detection of the Pyricularia oryzae lineage causing wheat blast. Plant Dis.

Urashima, A.S., Lavorent, N.A., Goulart, A.C.P., and Mehta, Y.R. (2004). Resistance spectra of wheat cultivars and virulence diversity of Magnaporthe grisea isolates in Brazil. Fitopatol. Bras. 29, 511–518.

Urashima, A.S., Leite, S.F., and Galbieri, R. (2007). Eficiência da disseminação aérea em Pyricularia grisea. Summa Phytopathol. 33, 275–279.

Yasuhara-Bell, J., Pedley, K.F., Farman, M., Valent, B., and Stack, J.P. (2018). Specific detection of the wheat blast pathogen (Magnaporthe oryzae Triticum) by loop-mediated isothermal amplification. Plant Dis.

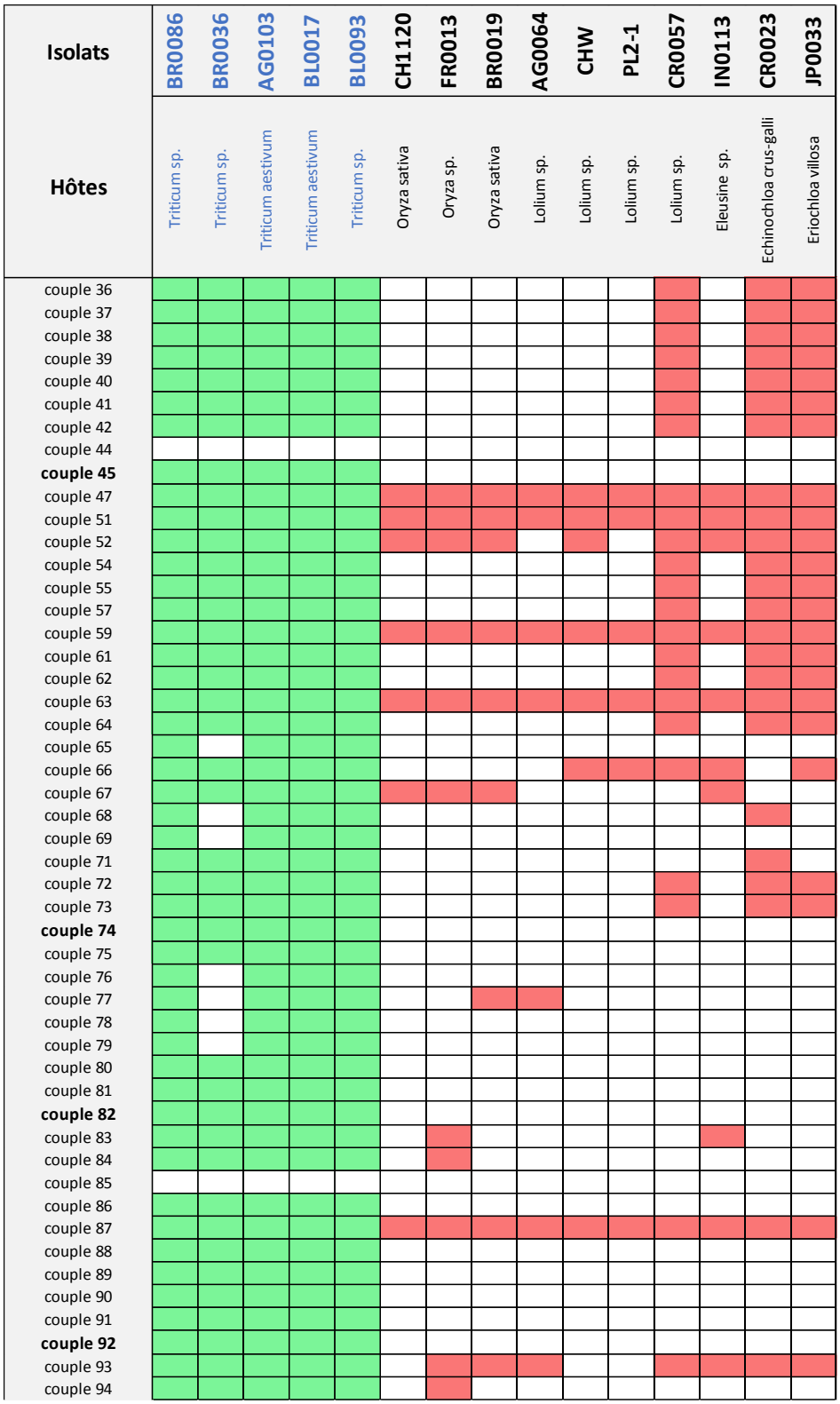
Isolates	Host of origin	PCR C74	PCR C82	PCR C92	PCR C45	qPCR C45	LAMP n°5	qPCR MoT3	qPCR C17
AG0103	Triticum aestivum	+	+	+	+	26,13 ± 0,16	3:47 ± 0:00	+	+
BL0017	Triticum aestivum	+	+	+	+	25,46 ± 0,03	3:48 ± 0:00	+	+
BL0018	Triticum aestivum	+	+	+	+	25,29 ± 0,05	3:51 ± 0:03	+	+
BL0020	Triticum aestivum	+	+	+	+	25,04 ± 0,05	3:53 ± 0:00	+	+
BL0023	Triticum aestivum	+	+	+	+	25,33 ± 0,10	3:54 ± 0:01	+	+
BL0028	Triticum aestivum	+	+	+	+	25,19 ± 0,20	3:53 ± 0:00	+	+
BL0037	Triticum aestivum	+	+	+	+	24,64 ± 0,06	3:55 ± 0:01	+	+
BL0044	Triticum aestivum	+	+	+	+	24,46 ± 0,19	3:55 ± 0:01	+	+
BL0046	Triticum aestivum	+	+	+	+	25,22 ± 0,20	3:52 ± 0:00	+	+
BL0063	Triticum aestivum	+	+	+	+	25,01 ± 0,05	3:54 ± 0:01	+	+
BL0092	Triticum sp.	+	+	+	+	23,75 ± 0,18	3:56 ± 0:01	+	+
BL0093	Triticum sp.	+	+	+	+	24,45 ± 0,20	3:52 ± 0:02	+	+
BR0031	Triticum sp.	+	-	+	+	24,67 ± 0,15	3:52 ± 0:02	+	+
BR0032	Triticum sp.	+	+	+	+	23,22 ± 0,30	3:57 ± 0:00	-	+
BR0034	Triticum sp.	-	+	+	+	24,43 ± 0,11	3:53 ± 0:01	+	+
BR0036	Triticum sp.	+	+	+	+	25,30 ± 0,13	3:48 ± 0:01	+	+
BR0039	Triticum sp.	-	+	+	+	24,61 ± 0,05	3:50 ± 0:02	+	+
BR0040	Triticum sp.	+	+	+	+	25,49 ± 0,10	3:51 ± 0:05	+	+
BR0041	Triticum sp.	-	+	+	+	26,30 ± 0,05	3:47 ± 0:03	+	+
BR0043	Triticum sp.	+	+	+	+	24,60 ± 0,13	3:55 ± 0:03	-	+
BR0045	Triticum sp.	+	+	+	+	27,79 ± 0,12	3:40 ± 0:03	+	+
BR0047	Triticum sp.	-	+	+	+	25,51 ± 0,06	3:53 ± 0:01	+	+
BR0080	Triticum sp.	+	+	+	+	25,09 ± 0,06	3:53 ± 0:00	+	+
BR0086	Triticum sp.	+	+	+	+	24,84 ± 0,21	3:51 ± 0:03	+	+
BR0087	Triticum sp.	+	+	+	+	25,76 ± 0,29	3:51 ± 0:03	+	+
BR0088	Triticum sp.	+	+	+	+	24,58 ± 0,05	3:54 ± 0:01	+	+
BR0123	Triticum aestivum	+	+	+	+	24,61 ± 0,07	3:55 ± 0:01	+	+
BTGP16	Triticum sp.	+	+	+	+	25,35 ± 0,18	3:70 ± 0:44	+	+
BTJP4-1	Triticum sp.	+	+	+	+	25,33 ± 0,07	3:38 ± 0:00	+	+
BTMP13-1	Triticum sp.	+	+	+	+	25,52 ± 0,25	3:37 ± 0:00	+	+
AG0102	Triticum aestivum	NT	NT	NT	-	-	-	NT	NT
BL0042	Triticum aestivum	NT	NT	NT	+	26,24 ± 0,09	3:48 ± 0:09	NT	NT
BL0066	Triticum aestivum	NT	NT	NT	+	25,55 ± 0,23	3:50 ± 0:01	NT	NT
BL0074	Triticum aestivum	NT	NT	NT	+	26,31 ± 0,08	3:41 ± 0:01	NT	NT
AG0054	Bromus sp.	+	+	+	-	-	10:39	-	-
AG0055	Bromus sp.	+	-	-	-	-	-	+	-
AG0061	Bromus unioloides	-	+	+	-	-	-	-	-
AG0062	Lolium sp.	-	-	-	-	-	-	-	-
AG0063	Lolium sp.	-	-	-	-	-	10:82 ± 0:52	-	-
AG0064	Lolium sp.	-	-	-	-	-	-	-	-
AG0065	Stenotaphrum sp.	-	-	-	-	-	-	-	-
AG0132	Oryza sativa	-	-	-	-	-	-	-	-
BF0017	Pennisetum typhoides	-	-	-	-	-	-	-	-
BR0019	Oryza sativa	-	-	-	-	-	-	-	-
BR0029	Digitaria sanguinalis	-	-	-	-	-	-	-	-
BR0030	Cenchrus echinatus	-	-	-	-	-	-	-	-
BR0062	Eleusine indica	-	-	-	-	-	-	-	-
BR0070	Eragrostis sp.	-	-	-	-	-	-	-	-
Br58	Avena sp.	-	-	-	-	37,04 ± 2,02	-	-	-
CD0143	Digitaria exilis	-	-	-	-	-	-	-	-
Cd88215	Echinochloa colona	-	-	-	-	-	3:65 ± 0:12	-	-
CH0333	Oryza sativa	-	-	-	-	-	-	-	-
CH1120	Oryza sativa	-	-	-	-	-	-	-	-
ch8401	?	-	-	-	-	37,49 ± 1,68	-	NT	NT
CHRF	Lolium sp.	-	-	-	-	-	9:66	-	-
CHW	Lolium sp.	-	-	-	-	-	-	-	-
CR0021	Panicum miliaceum	-	-	-	-	-	-	-	-
CR0023	Echinochloa crus-galli	-	-	-	-	-	3:52 ± 0:04	-	+
CR0026	Lolium sp.	-	-	-	-	-	-	-	-
CR0029	Festuca elalior	-	-	-	-	-	-	-	-
CR0030	Setaria viridis	-	-	-	-	-	-	-	-
CR0031	Setaria italica	-	-	-	-	-	-	-	-

Isolates	Host of origin	PCR C74	PCR C82	PCR C92	PCR C45	qPCR C45	LAMP n°5	qPCR MoT3	qPCR C17
CR0057	Lolium sp.	-	-	-	-	-	3:59 ± 0:02	-	+
EG0028	Cyperus rotundus	-	-	-	-	-	-	-	-
FH	Lolium sp.	-	-	-	-	-	-	-	-
FR0013	Oryza sp.	-	-	-	-	-	-	-	-
FR1069	Lolium sp.	-	-	-	-	-	-	-	-
GG11	Lolium sp.	-	-	-	-	-	-	-	-
GN0001	Zea mays	-	-	-	-	-	-	-	-
GR0001	Ctenanthe oppenheimiana	-	-	-	-	-	-	-	-
GY0011	Oryza sativa	-	-	-	-	-	-	-	-
HO	Lolium sp.	-	-	-	-	-	-	-	-
IN0003	Panicum repens	-	-	-	-	-	-	-	-
IN0005	Panicum maximun	+	-	-	-	-	-	-	-
IN0022	Setaria sp.	-	-	-	-	-	13:10	-	-
IN0023	Setaria sp.	-	-	-	-	-	-	-	-
IN0082	Oryza sativa	-	-	-	-	-	-	-	-
IN0108	Setaria sp.	-	-	-	-	-	-	-	-
IN0113	Eleusine sp.	-	-	-	-	-	-	-	-
IN0115	Oryza sativa	-	-	-	-	-	-	-	-
IR0013	Zea mays	-	-	-	-	-	-	-	-
IR0015	Zea mays	-	-	-	-	-	-	-	-
IR0095	Zea mays	-	-	-	-	-	-	-	-
IR0102	Echinochloa sp.	-	-	-	-	-	-	-	-
IS0001	Cyperus rotundus	-	-	-	-	-	-	-	-
JP0028	Eragrostis curvula	-	-	-	-	-	23:53	-	-
JP0030	Panicum bisulcatum	-	-	-	-	-	17:69	-	-
JP0031	Panicum coloratum	-	-	-	-	-	3:29 ± 0:50	-	-
JP0033	Eriochloa villosa	-	-	-	-	-	3:68 ± 0:04	-	+
JP0047	Hordeum vulgare	-	-	-	-	-	-	-	-
JP0048	Hordeum vulgare	-	-	-	-	-	-	-	-
KN0001	Hordeum vulgare	-	-	-	-	-	-	-	-
KN0006	Hordeum vulgare	-	-	-	-	-	-	-	-
Lc8401	Leptochloa chimensis	-	-	-	-	-	-	-	-
LpKY97	Lolium sp.	-	-	-	-	-	-	-	-
ML0031	Pennisetum sp.	-	-	-	-	-	-	-	-
Pd88413	Paspalum distichum	-	-	-	-	-	-	-	-
Pg1054	Stenotaphrum secundatum	-	-	-	-	-	-	-	-
Pg1213-22	Festuca sp.	-	-	-	-	-	-	-	-
PH0052	Cyperus rotundus	-	-	-	-	-	-	-	-
PH0053	Cyperus rotundus	-	-	-	-	-	-	-	-
PH0062	Paspalum distichum	-	-	-	-	-	-	-	-
PH0075	Brachiaria mutica	-	-	-	-	-	-	-	-
PH0078	Echinochloa sp.	-	-	-	-	-	-	-	-
PH0097	Paspalum paspaloides	-	-	-	-	-	-	-	-
PL 2-1	Lolium sp.	-	-	-	-	-	-	-	-
PL 3-1	Lolium sp.	-	-	-	-	-	-	-	-
PR0069	Stenotaphrum secundatum	-	-	-	-	-	-	-	-
Pr8202	Panicum repens	-	-	-	-	-	-	-	-
RW0043	Eleusine coracana	-	-	-	-	-	12:85	-	-
TF05-1	Festuca sp.	-	-	-	-	-	-	-	-
US0064	Setaria sp.	-	-	-	-	-	-	-	-
US0066	Cenchrus ciliaris	-	-	-	-	-	-	-	-
US0077	Lolium perenne	-	-	-	-	-	-	-	-
US0078	Lolium perenne	-	-	-	-	-	-	-	-
US0084	Stenotaphrum secundatum	-	-	-	-	-	-	-	-
VT0032	Leersia hexandra	-	-	-	-	-	-	-	-
AG0049	Echinochloa sp	NT	NT	NT	-	-	-	NT	NT
AG0050	Echinochloa sp	NT	NT	NT	-	-	-	NT	NT
AG0051	Echinochloa sp	NT	NT	NT	-	-	-	NT	NT
AG0058	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
AG0059	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
AG0067	Phalaris canariense	NT	NT	NT	-	-	3:68 ± 0:03	NT	NT
AU0002	Oryza rufipogon	NT	NT	NT	-	37,57 ± 1,30	8:51 ± 0:29	NT	NT

Isolates	Host of origin	PCR C74	PCR C82	PCR C92	PCR C45	qPCR C45	LAMP n°5	qPCR MoT3	qPCR C17
BF0026	Eleusine indica	NT	NT	NT	-	36,07 ± 1,68	-	NT	NT
BF0080	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
BF0083	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
BF0093	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
BF0181	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
BG0007	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
BG0023	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
BG0024	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
BR0066	Eleusine indica	NT	NT	NT	-	-	10:12 ± 0:88	NT	NT
BR0071	Echinochloa sp	NT	NT	NT	-	38,56 ± 0,54	-	NT	NT
BR0079	Eleusine indica	NT	NT	NT	-	38,48 ± 1,09	-	NT	NT
BR0093	Echinochloa colona	NT	NT	NT	-	-	-	NT	NT
CD0060	Oryza glaberrima	NT	NT	NT	-	-	15:62	NT	NT
CD0157	Eleusine indica	NT	NT	NT	-	37,80 ± 1,64	-	NT	NT
CD0258	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
CH0321	Oryza sativa	NT	NT	NT	-	-	-	NT	NT
CH0328	Oryza sativa	NT	NT	NT	-	-	-	NT	NT
CH0331	Oryza sativa	NT	NT	NT	-	-	-	NT	NT
CH0338	Oryza sativa	NT	NT	NT	-	-	-	NT	NT
CH0341	Oryza sativa	NT	NT	NT	-	-	-	NT	NT
CL0013	Echinochloa colona	NT	NT	NT	-	-	8:67 ± 0:04	NT	NT
CL0045	Rottboellia exalta	NT	NT	NT	-	-	11:45 ± 1:15	NT	NT
CL0089	Oryza rufipogon	NT	NT	NT	-	35,50 ± 2,39	-	NT	NT
CR0058	Setaria viridis	NT	NT	NT	-	-	-	NT	NT
CR0060	Eleusine indica	NT	NT	NT	-	-	11:96 ± 0:75	NT	NT
EG0025	Echinochloa colona	NT	NT	NT	-	-	-	NT	NT
FR1067	Lolium perenne	NT	NT	NT	-	-	21:75	NT	NT
GD0001	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
IN0004	Panicum repens	NT	NT	NT	-	-	-	NT	NT
IN0030	Echinochloa frumentaceum	NT	NT	NT	-	-	-	NT	NT
JP0020	Eleusine indica	NT	NT	NT	-	36,82 ± 2,22	-	NT	NT
JP0035	Pennisetum clandesti	NT	NT	NT	-	-	-	NT	NT
JP0036	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
JP0039	Anthoxanthum odoratum	NT	NT	NT	-	-	10:23	NT	NT
JP0040	Phalaris arundinacea	NT	NT	NT	-	-	-	NT	NT
JP0098	Setaria faberii	NT	NT	NT	-	-	21:96	NT	NT
JP0102	Setaria faberii	NT	NT	NT	-	-	-	NT	NT
MD0112	Eleusine indica	NT	NT	NT	-	-	11:91 ± 2:40	NT	NT
MD0153	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
ML0070	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
ML0074	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
NP0060	Eleusine coracana	NT	NT	NT	-	-	-	NT	NT
NR0041	Oryza longistaminata	NT	NT	NT	-	-	-	NT	NT
NR0049	Leersia hexandra	NT	NT	NT	-	-	-	NT	NT
PH0035	Brachiaria mutica	NT	NT	NT	-	-	18:42 ± 9:59	NT	NT
PH0045	Brachiaria mutica	NT	NT	NT	-	-	-	NT	NT
PH0046	Brachiaria distachya	NT	NT	NT	-	-	-	NT	NT
PH0056	Echinochloa ciliaris	NT	NT	NT	-	-	11:81 ± 0:67	NT	NT
PH0057	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
PH0077	Echinochloa colona	NT	NT	NT	-	-	-	NT	NT
PH0079	Panicum repens	NT	NT	NT	-	-	-	NT	NT
PH0080	Panicum repens	NT	NT	NT	-	-	-	NT	NT
PH0081	Paspalum paspaloides	NT	NT	NT	-	-	12:19 ± 2:68	NT	NT
PR0083	Stenotaphrum secondatum	NT	NT	NT	-	-	-	NT	NT
RN0001	Zingiber officinale	NT	NT	NT	-	-	-	NT	NT
RW0018	Eleusine coracana	NT	NT	NT	-	37,29 ± 2,65	13:56	NT	NT
RW0022	Eleusine indica	NT	NT	NT	-	-	-	NT	NT
RW0031	Eleusine coracana	NT	NT	NT	-	-	-	NT	NT
RW0036	Eleusine coracana	NT	NT	NT	-	-	-	NT	NT
RW0038	Eleusine coracana	NT	NT	NT	-	-	-	NT	NT
RW0041	Eleusine coracana	NT	NT	NT	-	-	25:09	NT	NT
LSV M 641	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT

Isolates	Host of origin	PCR C74	PCR C82	PCR C92	PCR C45	qPCR C45	LAMP n°5	qPCR MoT3	qPCR C17
LSV M 723	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 860	Blé	NT	NT	NT	-	-	NT	NT	NT
LSV M 702	Zae mays sp	NT	NT	NT	-	-	NT	NT	NT
LSV M 706	Zae mays sp	NT	NT	NT	-	-	NT	NT	NT
LNPV 269	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSVM 861	Blé	NT	NT	NT	-	-	NT	NT	NT
LSV M 273	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 811	Blé	NT	NT	NT	-	-	NT	NT	NT
LSV M 813	Blé	NT	NT	NT	-	-	NT	NT	NT
LSV M 662	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 694	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 697	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 642	Triticum sp.	NT	NT	NT	-	-	NT	NT	NT
LSV M 859	Blé	NT	NT	NT	-	-	NT	NT	NT
LSV M 863	Blé	NT	NT	NT	-	-	NT	NT	NT
CBS918.96	Dianthus chinonsis	NT	NT	NT	-	-	NT	NT	NT
CBS965.95	Triticum	NT	NT	NT	-	-	NT	NT	NT
BRIP46550	Malus	NT	NT	NT	-	-	NT	NT	NT

Figure 1: PCR, qPCR and LAMP primers screening using a large DNA panel. qPCR tests MoT3 (Pieck et al., 2017) and C17 (Thierry et al., 2019) are included for comparison. + : amplification; - : no amplification; NT: not tested. Units are Ct values for qPCR and minutes for LAMP.



Supplementary figure 1: PCR primers screening using a small panel. Green: Specific amplification; red: non-specific amplification; white: no amplification. Primers further studied are indicated in bold.

Isolate	BR0036	BR0031	BR0088	JP0033	AG0061	IN0113	CR0057
Host of origin	Trticum sp.	Trticum sp.	Trticum sp.	Eriochloa vilosa	Bromus unioloides	Eleusine sp.	Lolium sp.
LAMP n°1	5:07	8:79	8:03	19:38	12:38	12:33	13:40
LAMP n°2	-	-	-	-	-	-	-
LAMP n°3	-	-	-	-	-	-	-
LAMP n°4	19:74	19:51	NT	NT	20:84	21:60	NT
LAMP n°5	3:00	3:09	3:19	3:39	-	-	3:31

Supplementary figure 2: LAMP primers screening using a small panel. Green: specific amplification; red: non-specific amplification (minutes).

Supplementary table 1: List of DNAs used in this study

Isolats	Espèces	Hôtes	Année	Lieux	Origine
AG0049	Pyricularia oryzae	Echinochloa sp	2002	Argentine	UMR BGPI
AG0050	Pyricularia oryzae	Echinochloa sp	2001	Argentine	UMR BGPI
AG0051	Pyricularia oryzae	Echinochloa sp	2001	Argentine	UMR BGPI
AG0054	Pyricularia oryzae	Bromus sp.	2002	Argentine	UMR BGPI
AG0055	Pyricularia oryzae	Bromus sp.	2002	Argentine	UMR BGPI
AG0058	Pyricularia oryzae	Eleusine indica	2003	Argentine	UMR BGPI
AG0059	Pyricularia oryzae	Eleusine indica	2003	Argentine	UMR BGPI
AG0061	Pyricularia oryzae	Bromus unioloides	2003	Argentine	UMR BGPI
AG0062	Pyricularia oryzae	Lolium sp.	2004	Argentine	UMR BGPI
AG0063	Pyricularia oryzae	Lolium sp.	2004	Argentine	UMR BGPI
AG0064	Pyricularia oryzae	Lolium sp.	2004	Argentine	UMR BGPI
AG0065	Pyricularia oryzae	Stenotaphrum sp.	2007	Argentine	UMR BGPI
AG0067	Pyricularia oryzae	Phalaris canariense	2008	Argentine	UMR BGPI
AG0102	Pyricularia oryzae	Triticum aestivum	2015	Argentine	UMR BGPI
AG0103	Pyricularia oryzae	Triticum aestivum	2016	Argentine	UMR BGPI
AG0132	Pyricularia oryzae	Oryza sativa	2001	Argentine	UMR BGPI
AU0002	Pyricularia oryzae	Oryza rufipogon	1982	Australie	UMR BGPI
BF0017	Pyricularia oryzae	Pennisetum typhoides	1990	Burkina Faso	UMR BGPI
BF0026	Pyricularia oryzae	Eleusine indica	1990	Burkina Faso	UMR BGPI
BF0080	Pyricularia oryzae	Oryza longistaminata	2015	Burkina Faso	UMR BGPI
BF0083	Pyricularia oryzae	Oryza longistaminata	2015	Burkina Faso	UMR BGPI
BF0093	Pyricularia oryzae	Oryza longistaminata	2015	Burkina Faso	UMR BGPI
BF0181	Pyricularia oryzae	Oryza longistaminata	2014	Burkina Faso	UMR BGPI
BG0007	Pyricularia oryzae	Leersia hexandra	1989	Bangladesh	UMR BGPI
BG0023	Pyricularia oryzae	Leersia hexandra	1991	Bangladesh	UMR BGPI
BG0024	Pyricularia oryzae	Leersia hexandra	1992	Bangladesh	UMR BGPI
BL0017	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0018	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0020	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0023	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0028	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0037	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0042	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0044	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0046	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0063	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0066	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0074	Pyricularia oryzae	Triticum aestivum	2010	Bolivie	UMR BGPI
BL0092	Pyricularia oryzae	Triticum sp.			UMR BGPI
BL0093	Pyricularia oryzae	Triticum sp.			UMR BGPI
BR0019	Pyricularia oryzae	Oryza sativa	1986	Brésil	UMR BGPI
BR0029	Pyricularia oryzae	Digitaria sanguinalis	1989	Brésil	UMR BGPI
BR0030	Pyricularia oryzae	Cenchrus echinatus	1989	Brésil	UMR BGPI
BR0031	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0032	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0034	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0036	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0039	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0040	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0041	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0043	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0045	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0047	Pyricularia oryzae	Triticum sp.	1989	Brésil	UMR BGPI
BR0062	Pyricularia oryzae	Eleusine indica	1990	Brésil	UMR BGPI
BR0066	Pyricularia oryzae	Eleusine indica	1991	Brésil	UMR BGPI
BR0070	Pyricularia oryzae	Eragrostis sp.	1991	Brésil	UMR BGPI
BR0071	Pyricularia oryzae	Echinochloa sp	1990	Brésil	UMR BGPI
BR0079	Pyricularia oryzae	Eleusine indica	1991	Brésil	UMR BGPI
BR0080	Pyricularia oryzae	Triticum sp.		Brésil	UMR BGPI
BR0086	Pyricularia oryzae	Triticum sp.	1990	Brésil	UMR BGPI
BR0087	Pyricularia oryzae	Triticum sp.	1990	Brésil	UMR BGPI
BR0088	Pyricularia oryzae	Triticum sp.	1990	Brésil	UMR BGPI
BR0093	Pyricularia oryzae	Echinochloa colona	1990	Brésil	UMR BGPI
BR0123	Pyricularia oryzae	Triticum aestivum	1998	Brésil	UMR BGPI
Br58	Pyricularia oryzae	Avena sp.			Iwate Biotechnology Research Center
BTGP16	Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory
BTJP4-1	Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory
BTMP13-1	Pyricularia oryzae	Triticum sp.			The Sainsbury Laboratory
CD0060	Pyricularia oryzae	Oryza glaberrima	1983	Côte d'Ivoire	UMR BGPI
CD0143	Pyricularia oryzae	Digitaria exilis	1989	Côte d'Ivoire	UMR BGPI
CD0157	Pyricularia oryzae	Eleusine indica	1989	Côte d'Ivoire	UMR BGPI
CD0258	Pyricularia oryzae	Leersia hexandra	2014	Côte d'Ivoire	UMR BGPI
Cd88215	Pyricularia oryzae	Echinochloa colona			Department of plant pathology, university of Kentucky
CH0321	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH0328	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH0331	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH0333	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH0338	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH0341	Pyricularia oryzae	Oryza sativa	1998	Chine	UMR BGPI
CH1120	Pyricularia oryzae	Oryza sativa	2009	Chine	UMR BGPI
ch8401	Pyricularia oryzae	?			Department of plant pathology, university of Kentucky
CHRF	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
CHW	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
CL0013	Pyricularia oryzae	Echinochloa colona	1986	Colombie	UMR BGPI
CL0045	Pyricularia oryzae	Rottboellia exalta		Colombie	UMR BGPI
CL0089	Pyricularia oryzae	Oryza rufipogon	2001	Colombie	UMR BGPI
CR0021	Pyricularia oryzae	Panicum miliaceum		République de Corée	UMR BGPI
CR0023	Pyricularia oryzae	Echinochloa crus-galli		République de Corée	UMR BGPI
CR0026	Pyricularia oryzae	Lolium sp.		République de Corée	UMR BGPI
CR0029	Pyricularia oryzae	Festuca elalior		République de Corée	UMR BGPI
CR0030	Pyricularia oryzae	Setaria viridis	1991	République de Corée	UMR BGPI
CR0031	Pyricularia oryzae	Setaria italica	1991	République de Corée	UMR BGPI
CR0057	Pyricularia oryzae	Lolium sp.		République de Corée	UMR BGPI
CR0058	Pyricularia oryzae	Setaria viridis		République de Corée	UMR BGPI
CR0060	Pyricularia oryzae	Eleusine indica		République de Corée	UMR BGPI
EG0025	Pyricularia oryzae	Echinochloa colona		Egypte	UMR BGPI
EG0028	Pyricularia oryzae	Cyperus rotundus		Egypte	UMR BGPI
FH	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
FR0013	Pyricularia oryzae	Oryza sp.	1988	France	UMR BGPI
FR1067	Pyricularia oryzae	Lolium perenne	2017	France	UMR BGPI
FR1069	Pyricularia oryzae	Lolium sp.	2017	France	UMR BGPI

Isolats	Espèces	Hôtes	Année	Lieux	Origine
GD0001	Pyricularia oryzae	Eleusine indica	1991	Guadeloupe	UMR BGPI
GG11	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
GN0001	Pyricularia oryzae	Zea mays	1985	Gabon	UMR BGPI
GR0001	Pyricularia oryzae	Ctenanthe oppenheimiana	1998	Grèce	UMR BGPI
GY0011	Pyricularia oryzae	Oryza sativa	1978	Guyane	UMR BGPI
HO	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
IN0003	Pyricularia oryzae	Panicum repens		Inde	UMR BGPI
IN0004	Pyricularia oryzae	Panicum repens		Inde	UMR BGPI
IN0005	Pyricularia oryzae	Panicum maximun		Inde	UMR BGPI
IN0022	Pyricularia oryzae	Setaria sp.	1992	Inde	UMR BGPI
IN0023	Pyricularia oryzae	Setaria sp.	1992	Inde	UMR BGPI
IN0030	Pyricularia oryzae	Echinochloa frumentaceum	1993	Inde	UMR BGPI
IN0082	Pyricularia oryzae	Oryza sativa	1992	Inde	UMR BGPI
IN0108	Pyricularia oryzae	Setaria sp.	1992	Inde	UMR BGPI
IN0113	Pyricularia oryzae	Eleusine sp.	1992	Inde	UMR BGPI
IN0115	Pyricularia oryzae	Oryza sativa	1997	Inde	UMR BGPI
IR0013	Pyricularia oryzae	Zea mays	2012	Iran	UMR BGPI
IR0015	Pyricularia oryzae	Zea mays	2012	Iran	UMR BGPI
IR0095	Pyricularia oryzae	Zea mays	2016	Iran	UMR BGPI
IR0102	Pyricularia oryzae	Echinochloa sp.	2016	Iran	UMR BGPI
IS0001	Pyricularia oryzae	Cyperus rotundus	1979	Israël	UMR BGPI
JP0020	Pyricularia oryzae	Eleusine indica		Japon	UMR BGPI
JP0028	Pyricularia oryzae	Eragrostis curvula	1976	Japon	UMR BGPI
JP0030	Pyricularia oryzae	Panicum bisulcatum	1991	Japon	UMR BGPI
JP0031	Pyricularia oryzae	Panicum coloratum	1991	Japon	UMR BGPI
JP0033	Pyricularia oryzae	Eriochloa villosa		Japon	UMR BGPI
JP0035	Pyricularia oryzae	Pennisetum clandesti		Japon	UMR BGPI
JP0036	Pyricularia oryzae	Leersia hexandra		Japon	UMR BGPI
JP0039	Pyricularia oryzae	Anthoxanthum odoratum		Japon	UMR BGPI
JP0040	Pyricularia oryzae	Phalaris arundinacea		Japon	UMR BGPI
JP0047	Pyricularia oryzae	Hordeum vulgare	1980	Japon	UMR BGPI
JP0048	Pyricularia oryzae	Hordeum vulgare	1980	Japon	UMR BGPI
JP0098	Pyricularia oryzae	Setaria faberii	2005	Japon	UMR BGPI
JP0102	Pyricularia oryzae	Setaria faberii	2005	Japon	UMR BGPI
KN0001	Pyricularia oryzae	Hordeum vulgare	1994	Kenya	UMR BGPI
KN0006	Pyricularia oryzae	Hordeum vulgare	1994	Kenya	UMR BGPI
Lc8401	Pyricularia oryzae	Leptochloa chimensis			Department of plant pathology, university of Kentucky
LpKY97	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
MD0112	Pyricularia oryzae	Eleusine indica	1990	Madagascar	UMR BGPI
MD0153	Pyricularia oryzae	Eleusine indica	2001	Madagascar	UMR BGPI
ML0031	Pyricularia oryzae	Pennisetum sp.	1990	Mali	UMR BGPI
ML0070	Pyricularia oryzae	Oryza longistaminata	2013	Mali	UMR BGPI
ML0074	Pyricularia oryzae	Oryza longistaminata	2013	Mali	UMR BGPI
NP0060	Pyricularia oryzae	Eleusine coracana	2008	Nepal	UMR BGPI
NR0041	Pyricularia oryzae	Oryza longistaminata	2013	Niger	UMR BGPI
NR0049	Pyricularia oryzae	Leersia hexandra	2013	Niger	UMR BGPI
Pd88413	Pyricularia oryzae	Paspalum distichum			Department of plant pathology, university of Kentucky
Pg1054	Pyricularia oryzae	Stenotaphrum secundatum			Department of plant pathology, university of Kentucky
Pg1213-22	Pyricularia oryzae	Festuca sp.			Department of plant pathology, university of Kentucky
PH0035	Pyricularia oryzae	Brachiaria mutica	1983	Philippines	UMR BGPI
PH0045	Pyricularia oryzae	Brachiaria mutica	1990	Philippines	UMR BGPI
PH0046	Pyricularia oryzae	Brachiaria distachya	1990	Philippines	UMR BGPI
PH0052	Pyricularia oryzae	Cyperus rotundus	1990	Philippines	UMR BGPI
PH0053	Pyricularia oryzae	Cyperus rotundus	1990	Philippines	UMR BGPI
PH0056	Pyricularia oryzae	Echinochloa ciliaris	1990	Philippines	UMR BGPI
PH0057	Pyricularia oryzae	Eleusine indica	1990	Philippines	UMR BGPI
PH0062	Pyricularia oryzae	Paspalum distichum	1990	Philippines	UMR BGPI
PH0075	Pyricularia oryzae	Brachiaria mutica	1989	Philippines	UMR BGPI
PH0077	Pyricularia oryzae	Echinochloa colona	1983	Philippines	UMR BGPI
PH0078	Pyricularia oryzae	Echinochloa sp.	1989	Philippines	UMR BGPI
PH0079	Pyricularia oryzae	Panicum repens	1983	Philippines	UMR BGPI
PH0080	Pyricularia oryzae	Panicum repens	1983	Philippines	UMR BGPI
PH0081	Pyricularia oryzae	Paspalum paspaloides	1983	Philippines	UMR BGPI
PH0097	Pyricularia oryzae	Paspalum paspaloides	1983	Philippines	UMR BGPI
PL 2-1	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
PL 3-1	Pyricularia oryzae	Lolium sp.			Department of plant pathology, university of Kentucky
PR0069	Pyricularia oryzae	Stenotaphrum secundatum	1992	Portugal	UMR BGPI
PR0083	Pyricularia oryzae	Stenotaphrum secondatum	1992	Portugal	UMR BGPI
Pr8202	Pyricularia oryzae	Panicum repens			Department of plant pathology, university of Kentucky
RN0001	Pyricularia oryzae	Zingiber officinale		Ile de la reunion	UMR BGPI
RW0018	Pyricularia oryzae	Eleusine coracana	1990	Rwanda	UMR BGPI
RW0022	Pyricularia oryzae	Eleusine indica	1990	Rwanda	UMR BGPI
RW0031	Pyricularia oryzae	Eleusine coracana	1990	Rwanda	UMR BGPI
RW0036	Pyricularia oryzae	Eleusine coracana	1990	Rwanda	UMR BGPI
RW0038	Pyricularia oryzae	Eleusine coracana	1990	Rwanda	UMR BGPI
RW0041	Pyricularia oryzae	Eleusine coracana	1990	Rwanda	UMR BGPI
RW0043	Pyricularia oryzae	Eleusine coracana		Rwanda	UMR BGPI
TF05-1	Pyricularia oryzae	Festuca sp.			Department of plant pathology, university of Kentucky
US0064	Pyricularia oryzae	Setaria sp.	1991	USA	UMR BGPI
US0066	Pyricularia oryzae	Cenchrus ciliaris	1995	USA	UMR BGPI
US0077	Pyricularia oryzae	Lolium perenne		USA	UMR BGPI
US0078	Pyricularia oryzae	Lolium perenne	2000	USA	UMR BGPI
US0084	Pyricularia oryzae	Stenotaphrum secundatum	2000	USA	UMR BGPI
VT0032	Pyricularia oryzae	Leersia hexandra	2002	Vietnam	UMR BGPI
LSV M 641	Microdochium nivale	Triticum sp.			LSV
LSV M 723	Fusarium tricinctum	Triticum sp.			LSV
LSV M 860	Fusarium tricinctum	Blé			LSV
LSV M 702	Fusarium proliferatum	Zae mays sp			LSV
LSV M 706	Fusarium proliferatum	Zae mays sp			LSV
LNPV 269	Fusarium poae	Triticum sp.			LSV
LSVM 861	Fusarium poae	Blé			LSV
LSV M 273	Fusarium graminearum	Triticum sp.			LSV
LSV M 811	Fusarium graminearum	Blé			LSV
LSV M 813	Fusarium graminearum	Blé			LSV
LSV M 662	Fusarium culmorum	Triticum sp.			LSV
LSV M 694	Fusarium culmorum	Triticum sp.			LSV
LSV M 697	Fusarium culmorum	Triticum sp.			LSV
LSV M 642	Fusarium avenaceum	Triticum sp.			LSV
LSV M 859	Fusarium avenaceum	Blé			LSV
LSV M 863	Fusarium avenaceum	Blé			LSV
CBS918.96	Alternaria tenuissima	Dianthus chinonsis			LSV
CBS965.95	Alternaria tenuissima	Triticum			LSV
BRIP46550	Alternaria alternata	Malus			LSV

Supplementary table 2: List of primer used in this study

Test names	Primer types	scaffold BR0032	start fragment	Prime names	Sequence 5'-3'	Size	GC%	Hairpin Tm	Tm homodimèr	Tm	Size
C36	forward	15	373000	15_373000_3f	GCTACGCATTTTTATTCATATTTTAAACC	29	27.6	None	None	59.6	147pb
	probe	15	373000	15_373000_5p	GTCTAGACCTAGTAGCTACCTCATGGATTTTGC	33	45.5	34.9	24.2	67.7	
	reverse	15	373000	15_373000_4r	GCCGACTCGATACAGATATTC	21	47.6	None	2.7	58.9	
C37	forward	15	373000	15_373000_3f	GCTACGCATTTTTATTCATATTTTAAACC	29	27.6	None	None	59.6	140pb
	probe	15	373000	15_373000_5p	GTCTAGACCTAGTAGCTACCTCATGGATTTTGC	33	45.5	34.9	24.2	67.7	
	reverse	15	373000	15_373000_1_2A_r	TCGATACAGATATTCGGTTCGTC	23	43.5	32.8	None	60.8	
C38	forward	15	373000	15_373000_4f	GATTTTGCTTCTCTCGGTTTGA	22	40.9	None	None	60.2	157pb
	probe	15	373000	15_373000_6p	GGAAGATACTGGAATTAAACAGCTTCGCC	29	44.8	38.6	None	65.6	
	reverse	15	373000	15_373000_5r	CTCCTCACTTATCGTCAGTGTAG	23	47.8	42.6	13.5	60.7	
C39	forward	15	373000	15_373000_6f	CCGAATATCTGTATCGAGTCGG	22	50	None	18.9	60.9	131pb
	probe	15	373000	15_373000_6p	GGAAGATACTGGAATTAAACAGCTTCGCC	29	44.8	38.6	None	65.6	
	reverse	15	373000	15_373000_5r	CTCCTCACTTATCGTCAGTGTAG	23	47.8	42.6	13.5	60.7	
C40	forward	15	373000	15_373000_5f	AGTTTAACTAGGAAACAAGCTACG	24	37.5	39.4	22.5	59.8	151pb
	probe	15	373000	15_373000_6p	GGAAGATACTGGAATTAAACAGCTTCGCC	29	44.8	38.6	None	65.6	
	reverse	15	373000	15_373000_6r	GATATTCGGTTCGTCAAACCG	21	47.6	67.6	24.5	60.4	
C41	forward	15	373000	15_373000_1f	CATTCTCGGTTTTGCATTAAGTCTAGAC	28	39.3	None	15.4	60.9	89pb
	probe	15	373000	15_373000_1_2_p	AGCTACCTCATGGATTTTGCTTCTCTCGG	29	48.3	37	None	66.4	
	reverse	15	373000	15_373000_1_2_r	ACTCGATACAGATATTCGGTTCGTC	25	44	32.3	None	60.3	
C42	forward	15	373000	15_373000_1f	CATTCTCGGTTTTGCATTAAGTCTAGAC	28	39.3	None	15.4	60.9	93pb
	probe	15	373000	15_373000_1_2_p	AGCTACCTCATGGATTTTGCTTCTCTCGG	29	48.3	37	None	66.4	
	reverse	15	373000	15_373000_4r	GCCGACTCGATACAGATATTC	21	47.6	None	2.7	58.9	
C44	forward	10	1275750	10_1275750_1f	AGGATGGCGCACCACTG	17	64.7	53.3	13.3	61.5	106pb
	probe	10	1275750	10_1275750_1p	TCGGGAAGAGGCGGGGCG	18	77.8	None	None	68.6	
	reverse	10	1275750	10_1275750_3r	CCGCTCAAGTATCTCGGTCTC	21	57.1	36.1	None	61.6	
C45	forward	15	720750	15_720750_1f	TCTTTCACTCCTCCGAAAGAC	21	47.6	48.1	14	60.4	102pb
	probe	15	720750	15_720750_1p	TGCCCTCATCAAAACCTGCAGCCAT	25	52	None	11.2	68.8	
	reverse	15	720750	15_720750_1r	GTATAGCTGGGTATCTTGGTAGAC	24	45.8	32.5	None	60.6	
C47	forward	15	401250	15_401250_1f	ACGCATTCAACAAACTCTTGAT	22	36.4	36.7	None	60.1	187pb
	probe	15	401250	15_401250_1p	CACTAACCAAGCTCATGGTTGCGGGAG	27	55.6	56.6	28.3	69.1	
	reverse	15	401250	15_401250_1r	GCCAGTCGTCACATCATATG	20	50	None	None	59.5	
C51	forward	15	402250	15_402250_1f	TGTGTGTTAGATGAGAGGCTG	21	47.6	None	None	60.5	111pb
	probe	15	402250	15_402250_1p	CGATGGTTTGGAAGATTTCGAAACAGGTCG	30	46.7	51.2	23.7	67.8	
	reverse	15	402250	15_402250_1r	CACTATGGAGGAAATCAACATAAAAAGA	28	32.1	None	None	60.9	
C52	forward	15	402250	15_402250_1f	TGTGTGTTAGATGAGAGGCTG	21	47.6	None	None	60.5	186pb
	probe	15	402250	15_402250_1p	CGATGGTTTGGAAGATTTCGAAACAGGTCG	30	46.7	51.2	23.7	67.8	
	reverse	15	402250	15_402250_2r	CCCCTTTTTTTGTCGTCGTT	20	45	None	None	60.3	
C54	forward	15	402250	15_402250_3f	GTGTGTTAGATGAGAGGCTGTG	22	50	None	None	61.7	114pb
	probe	15	402250	15_402250_1p	CGATGGTTTGGAAGATTTCGAAACAGGTCG	30	46.7	51.2	23.7	67.8	
	reverse	15	402250	15_402250_2r	CCCCTTTTTTTGTCGTCGTT	20	45	None	None	60.3	
C55	forward	15	426750	15_426750_1f	ATCCTTACAATGTTGTCTTCGATG	24	37.5	42.7	6.2	60.1	114pb
	probe	15	426750	15_426750_1p	ATCGTGCCATCTGACGTCTATGGTGC	26	53.8	42.3	34.1	68.3	
	reverse	15	426750	15_426750_1r	GCGCAAATTGTTGCCTCG	18	55.6	55.4	7.1	61.9	
C57	forward	15	374250	15_374250_1f	GGGCGAGTTCCAGGTC	16	68.8	None	None	60	102pb
	probe	15	374250	15_374250_1p	TGTCGACACAAGATTATTGTCGAACATGT	29	37.9	60.2	25.1	64.9	
	reverse	15	374250	15_374250_1r	CCTCATGATATAAGCCCATACATG	24	41.7	36.4	None	59.6	
C59	forward	15	405000	15_405000_1f	ATGATAGGGTTCTGTAGATCTAGAC	25	40	None	7.5	59.5	176pb
	probe	15	405000	15_405000_1p	GCCCTATCCACGAGTGCGTTAGCTTCAAC	29	55.2	52.4	None	70.4	
	reverse	15	405000	15_405000_1r	GATATTTATCTCGGATCATCGTGG	24	41.7	None	None	59.7	
C61	forward	15	403000	15_403000_1f	CGATGCACTGGATGAAGC	18	55.6	None	None	59.6	180pb
	probe	15	403000	15_403000_1p	GAACAGGAACAGGTCGTCAGTTGCCATTAA	30	46.7	35.1	None	68.4	
	reverse	15	403000	15_403000_2r	GGGGGTTAATGGTTTCTGTTTTC	23	43.5	None	None	61	
C62	forward	15	403000	15_403000_1f	CGATGCACTGGATGAAGC	18	55.6	None	None	59.6	136pb
	probe	15	403000	15_403000_1p	GAACAGGAACAGGTCGTCAGTTGCCATTAA	30	46.7	35.1	None	68.4	
	reverse	15	403000	15_403000_1r	CTTTCATGATTAGTTAGGATGGCA	24	37.5	None	None	59.6	
C63	forward	15	517000	15_517000_1f	GAAATAGGGAAACGTATGTGAACAG	25	40	40.9	None	60.9	213pb
	probe	15	517000	15_517000_1p	TAGACTTACCTGGACGCAATACGGTGATCC	30	50	49.9	0.4	68.6	
	reverse	15	517000	15_517000_1r	AGCTTCCTTTGGCTCGTTT	19	47.4	33.5	None	60.9	
C64	forward	15	584000	15_584000_1f	ATTTCGCCATCAGCGAGA	18	50	70.1	14.7	60.2	203pb
	probe	15	584000	15_584000_1p	AGGTCGTGTTCTTGGCGATCTCGCC	25	60	71.8	35.6	70.4	
	reverse	15	584000	15_584000_1r	ACCGACGTTGCCAGTATT	18	50	None	0	60	
C65	forward	17	431500	17_431500_1f	CCGGAATTTTGACTTGATGCAA	22	40.9	None	None	61.1	167pb
	probe	17	431500	17_431500_1p	TGGGGTCAGGATGCTTCACTGGGATTAACA	30	50	35.1	15.6	70.4	
	reverse	17	431500	17_431500_1r	TCAGGCGTTGCGCAA	15	60	56.5	26.9	60.6	
C66	forward	17	431750	17_431750_1f	ATCGTCACCATTACCCTCAC	20	50	None	None	60.2	135pb
	probe	17	431750	17_431750_1p	GAGCTGTAGTAGTACGATCTGGTATCCTCA	30	46.7	31.4	None	65.8	
	reverse	17	431750	17_431750_1r	ACCAACATGGGTGGGATC	18	55.6	41.7	22.7	60	
C67	forward	17	431750	17_431750_2f	TCCAAGCTGCATCGCG	16	65.5	34.6	2.5	61.3	113pb
	probe	17	431750	17_431750_2p	CATGTTGGTCAAGCTCTTGTGAGTCTAGTC	30	46.7	42.5	11	66.6	
	reverse	17	431750	17_431750_2r	CAGGCCGACTAGAGTCG	17	64.7	59.6	26.5	59.5	
C68	forward	17	431750	17_431750_3f	TTAATAATCCAGCAATCACTCGC	23	39.1	None	None	60.2	117pb
	probe	17	431750	17_431750_3p	CATCATCCTCGTTGTCCTCATCACCCTCG	29	55.2	None	None	69.5	
	reverse	17	431750	17_431750_3r	TACAGCTCTGACGAAGATGG	20	50	33.8	8.6	59.9	
C69	forward	17	432250	17_432250_1f	ACAGGTCAACAGGTGAAAAAC	21	42.9	38.7	None	59.9	155pb
	probe	17	432250	17_432250_1p	TGATGGTTGTGGAACCAGATCCCAAAAACA	30	43.3	60	14.2	69.4	
	reverse	17	432250	17_432250_1r	GGGCGGCGAGTTCATT	16	62.5	None	19.5	60.8	
C71	forward	15	372750	15_372750_3f	GGTTTTGCATTAAGTCTAGACCT	23	39.1	None	24.3	59.4	125pb
	probe	15	372750	15_372750_1p	GACGAACCGAATATCTGTATCGAGTCGGCG	30	53.3	40.6	15.8	69.5	
	reverse	15	372750	15_372750_1r	CGAAGCTGTTTAATTCCAGTATCTTC	26	38.5	None	None	60	
C72	forward	15	372750	15_372750_1f	CTCGGTTTTGCATTAAGTCTAGAC	24	41.7	None	20.6	60.7	128pb
	probe	15	372750	15_372750_1p	GACGAACCGAATATCTGTATCGAGTCGGCG	30	53.3	40.6	15.8	69.5	
	reverse	15	372750	15_372750_1r	CGAAGCTGTTTAATTCCAGTATCTTC	26	38.5	None	None	60	
C73	forward	15	372750	15_372750_2f	AATAGTTTAACTAGGAAACAAGCTACG	27	33.3	39.4	22.5	60.4	124pb
	probe	15	372750	15_372750_2p	GAAGATGCTACTTCCCTCTTTCATTCTCGG	30	46.7	40.4	12.4	66.4	
	reverse	15	372750	15_372750_2r	AAATCCATGAGGTAGCTACTAGG	23	43.5	36.1	27.4	59.9	

Test names	Primer types	scaffold BR0032	start fragment	Prime names	Sequence 5'-3'	Size	GC%	Hairpin Tm	Tm homodimèr	Tm	Size
	forward	17	24750	17_24750_1f	CCAAATACACCAGACCTAGTTCTATAA	27	37	None	None	37	
C74	probe	17	24750	17_24750_1p	GGCCAAGGGCATGGCTTTGTTGCT	24	58.3	59.4	34.4	70.3	175pb
	reverse	17	24750	17_24750_1r	GTTTTGCTGTTGGATGTTTACGAA	24	37.5	None	None	61.7	
	forward	17	24750	17_24750_2f	CCCAAATACACCAGACCTAGTT	22	45.5	None	None	60.5	
C75	probe	17	24750	17_24750_1p	GGCCAAGGGCATGGCTTTGTTGCT	24	58.3	59.4	34.4	70.3	176pb
	reverse	17	24750	17_24750_1r	GTTTTGCTGTTGGATGTTTACGAA	24	37.5	None	None	61.7	
	forward	17	186750	17_186750_1f	CTGCGGCTGCAACATTG	17	58.8	54	28.2	60.9	
C76	probe	17	186750	17_186750_2p	AGCGGCCTTGGGGACAGCAG	20	70	42.5	5.1	69.1	100pb
	reverse	17	186750	17_186750_1r	CCCAAGGCTGCTACCCT	17	64.7	None	3	61.9	
	forward	17	562750	17_562750_2f	CCAGCTTCCCGCACATAAG	19	57.9	42.9	None	61.9	
C77	probe	17	562750	17_562750_2p	GCAAACGAGCCAACAGATTTGGAAGGTCAC	30	50	42.5	9.3	69.8	196pb
	reverse	17	562750	17_562750_2r	ATGGTTCTGAAATAGCTCTGGAG	23	43.5	None	None	60.9	
	forward	17	562750	17_562750_2f	CCAGCTTCCCGCACATAAG	19	57.9	42.9	None	61.9	
C78	probe	17	562750	17_562750_2p	GCAAACGAGCCAACAGATTTGGAAGGTCAC	30	50	42.5	9.3	69.8	209pb
	reverse	17	562750	17_562750_3r	TGGGAATCCGTCCATGGT	18	55.6	50.5	18.2	61.4	
	forward	17	562750	17_562750_1f	CGTGGAAGTAAAAGCAAACGA	21	42.9	32.2	None	60.4	
C79	probe	17	562750	17_562750_1p	GAAGGTCACTTCTAACATACCAGCTGAGGC	30	50	48	19.9	68	147pb
	reverse	17	562750	17_562750_1r	CCATGGTTCTGAAATAGCTCTG	22	45.5	None	8.6	59.8	
	forward	44	105000	44_105000_1f	CCGGGGCCCAATTTTAGG	18	61.1	None	32.6	61.6	
C80	probe	44	105000	44_105000_1p	CGGAATTCCACGCAAAGGTACGGCTATTAG	30	50	35.4	26.1	68.8	87pb
	reverse	44	105000	44_105000_1r	GTGGGCCTTAACAGGCAT	18	55.6	65.3	19.7	60.8	
	forward	44	105000	44_105000_2f	ATGCCTGTTAAGGCCCAC	18	55.6	65.1	18.6	60.8	
C81	probe	44	105000	44_105000_2p	TTATCCGGGACGAGTGCGGTACGGAA	26	57.7	68.7	19.4	70.6	132pb
	reverse	44	105000	44_105000_2r	TACCGCTGCCTGTAGGAT	18	55.6	37.7	6.6	60.8	
	forward	44	105000	44_105000_2f	ATGCCTGTTAAGGCCCAC	18	55.6	65.1	18.6	60.8	
C82	probe	44	105000	44_105000_2p	TTATCCGGGACGAGTGCGGTACGGAA	26	57.7	68.7	19.4	70.6	228pb
	reverse	44	105000	44_105000_4r	TRTTGGAGGTAAATCCATGGTAAA	24	37.5	49	18.2	61	
	forward	44	105000	44_105000_3f	CGTATYTGGGACCGCCTA	18	55.6	None	None	59.9	
C83	probe	44	105000	44_105000_3p	GGGGCCCAATTTTAGGGYAACGGAATT	27	51.9	67.2	29.3	69.5	115pb
	reverse	44	105000	44_105000_3r	TTAACAGGCATTTGGTTGCAG	21	42.9	45	3	60.7	
	forward	44	105250	44_105250_1f	GCTYAGGAAATTYACGGCA	19	52.6	None	None	61.8	
C84	probe	44	105250	44_105250_1p	AGGAAGGTRAACGACGCATTRGTTGCAAAA	30	43.3	53.3	0.5	68.8	112pb
	reverse	44	105250	44_105250_1r	ACAYYTCGCTGCCCAAAAYTAATA	24	33.3	None	None	60.9	
	forward	44	105250	44_105250_2f	CTAYYAAYTATATTARTTTTGGGCAGC	27	37	41.5	19.3	61.5	
C85	probe	44	105250	44_105250_2p	TGGACCATATAAARTCGTTTGCGTAACCGA	30	43.3	39.7	3.7	67.7	116pb
	reverse	44	105250	44_105250_2r	GGGARCGYYATTGTAACGTTA	21	47.6	49.2	9	62	
	forward	44	105250	44_105250_3f	CCATATAAARTCGTTTGCGTAACC	24	41.7	39.7	None	61.2	
C86	probe	44	105250	44_105250_3p	GTTACAATRRCGYTCCCAAACGGCAATTTT	30	43.3	70.6	34.8	68.8	106pb
	reverse	44	105250	44_105250_3r	TTTGGTATTTGTAATACGGTTTTACGTA	28	28.6	51.2	17.7	60.6	
	forward	44	105250	44_105250_3f	CCATATAAARTCGTTTGCGTAACC	24	41.7	39.7	None	61.2	
C87	probe	44	105250	44_105250_3p	GTTACAATRRCGYTCCCAAACGGCAATTTT	30	43.3	70.6	34.8	68.8	112pb
	reverse	44	105250	44_105250_4r	GYTRTTTTTGGTATTTGTAATACGGT	26	34.6	None	None	61.6	
	forward	44	108250	44_108250_1f	GCCCGAAATTTGCCTGAAA	19	47.4	None	None	60.4	
C88	probe	44	108250	44_108250_1p	AGTCCGGACGAGGCCGGTARTAGG	24	62.5	63.3	34.8	69	106pb
	reverse	44	108250	44_108250_1r	CGCCCCCGTTCCG	13	84.6	None	None	60	
	forward	44	108250	44_108250_1f	GCCCGAAATTTGCCTGAAA	19	47.4	None	None	60.4	
C89	probe	44	108250	44_108250_1p	AGTCCGGACGAGGCCGGTARTAGG	24	62.5	63.3	34.8	69	89pb
	reverse	44	108250	44_108250_2r	CGTTGGCGCCTCCG	14	78.6	38.1	35.1	61.3	
	forward	44	108250	44_108250_3f	CCCGAAATTTGCCTGAAATAGG	22	45.5	None	50.4	60.9	
C90	probe	44	108250	44_108250_1p	AGTCCGGACGAGGCCGGTARTAGG	24	62.5	63.3	34.8	69	83pb
	reverse	44	108250	44_108250_3r	GCGCCTCCGCCTT	13	76.9	54.4	None	59.4	
	forward	44	108250	44_108250_4f	CGCGGGGGTTAAGGC	15	73.3	None	None	60.9	
C91	probe	44	108250	44_108250_2p	TCCGGAACGGGGGCGGC	17	82.4	75.7	1.7	69.3	132pb
	reverse	44	108250	44_108250_4r	ATTCGTTAAATATATCGTTATTAACGCAAA	30	23.3	53.5	26	60	
	forward	44	108250	44_108250_5f	AGCCGCGGGGGTTA	14	71.4	43.4	34.9	60.7	
C92	probe	44	108250	44_108250_2p	TCCGGAACGGGGGCGGC	17	82.4	75.7	1.7	69.3	135pb
	reverse	44	108250	44_108250_4r	ATTCGTTAAATATATCGTTATTAACGCAAA	30	23.3	53.5	26	60	
	forward	44	427750	44_427750_1f	GGTTCCAAACGTCATGAATATGT	23	39.1	36.1	None	60.3	
C93	probe	44	427750	44_427750_1p	CTCTTCAGTGCGACACCCCACGG	23	65.2	37.4	10.4	70.3	162pb
	reverse	44	427750	44_427750_1r	AGCCCTTCATGCTCCTTAC	19	52.6	33.1	None	60.2	
	forward	44	101750	44_101750_1f	GTGGAGARATCACGTTGCA	19	47.4	None	None	59.4	
C94	probe	44	101750	44_101750_1p	CGTRATCTCGGAACACGTCGCAGAATCT	28	50	44.9	None	68.1	83pb
	reverse	44	101750	44_101750_1r	AACTATYTACGCGGTTGCATA	21	42.9	43.3	18.5	60.4	
	F3	15	720750	15_720750_F3_1	GATATGCGCCGAAGTTCA						
LAMP n°1	B3	15	720750	15_720750_B3_1	GGCTAGAATAAGGTAAGAAGGAA						
	FIP	15	720750	15_720750_FIP_1	GGAGGAGTGAAAGACAGAAAAAGAATCTCTT						
	BIP	15	720750	15_720750_BIP_1	CACGTTCTGCTCTTGCCCTCGCTGGGTATCTTGG						
	F3	44	108250	44_108250_F3_1	GTTAATAACGATATATYTRACGART						
LAMP n°2	B3	44	108250	44_108250_B3_1	TTATGTAAACGGTRAATRTGTT						
	FIP	44	108250	44_108250_FIP_1	YTYTTCCTTYTAAYTTTCCGAYTYAGTTGGGAGT						
	BIP	44	108250	44_108250_BIP_1	AARTCCGGAAAAAAAGCGCTTTTATRAATRTTTA						
	F3	44	108250	44_108250_F3_2	GCGTTAATAACGATATATYTRACG						
LAMP n°3	B3	44	108250	44_108250_B3_2	TTATGTAAACGGTYAATYTGTT						
	FIP	44	108250	44_108250_FIP_2	AYTYTTCCTTYTAAYTTTCCGAYTYGTTGGGAGT						
	BIP	44	108250	44_108250_BIP_2	TAAARTCCGGAAAAAAGCGCTTTTARAATRTTT						
	F3	15	720750	15_720750_F3_2	TTCACTCCTCCGAAAGAC						
LAMP n°4	B3	15	720750	15_720750_B3_2	ATCGTCGTGGATCTTGCT						
	FIP	15	720750	15_720750_FIP_2	GCTGGGTATCTTGGTAGACTATACTTGCCCTCA						
	BIP	15	720750	15_720750_BIP_2	TTACCTTATTCTAGCCATCACATCCCAAAAGAA						
	F3	15	42800	15_42800_F3_2	TGCGTGATCAACGAATGGC						
LAMP n°5	B3	15	42800	15_42800_B3_2	CGGAAGCAAACTCTGCGATT						
	FIP	15	42800	15_42800_FIP_2	GCAAGATGCCTACCGTGGGGTTTTCTGGGTTCC						
	BIP	15	42800	15_42800_BIP_2	TCGTAAGGAGCATGAAGGGCTGTTTTGCGGAC						

CHAPITRE 2

-

Structure des populations au sein de la lignée Oryza

1- Contexte de l'étude

La structure des populations à l'échelle mondiale des isolats de *P. oryzae* pathogènes du riz a permis l'identification de groupes génétiques majeurs au sein de la lignée Oryza. Toutefois, les études ayant analysé cette structure ont inclus soit un grand nombre d'isolats mais génotypés avec un faible nombre de marqueurs génétiques (Saleh et al., 2014; Tharreau et al., 2009), soit au contraire un nombre important de marqueurs mais sur un nombre d'isolats réduits (Gladieux et al., 2018a; Zhong et al., 2018) et n'ont pas couvert les mêmes zones géographiques. L'Amérique du sud et l'Afrique restent d'ailleurs très peu étudiées. Ces différences ont entrainé des conclusions légèrement différentes entre les différentes études notamment en termes de nombre de lignées identifiées au sein de la lignée hôte-spécifique Oryza (de trois à six lignées en fonction de l'étude). Ces études ont pu mettre en évidence certaines différenciations entre les lignées en termes de mode de reproduction (reproduction sexuée ou asexuée selon la lignée) ou de largeur de spectre d'hôte (virulence ou avirulence sur des variétés de riz de type Indica) donnant quelques pistes sur les facteurs pouvant entrer en jeu dans la structuration de ces populations (Gladieux et al., 2018a; Saleh et al., 2014; Zhong et al., 2018). La structure à l'échelle globale de la lignée Oryza, les aires de répartition des lignées, ainsi que les facteurs impliqués dans le maintien de cette structure génétique restent donc à affiner.

Ce chapitre est présenté sous la forme d'un article en préparation.

Article 4: Barriers to gene flow between pandemic lineages of the rice blast fungus *Pyricularia oryzae*

Les principaux objectifs de cet article étaient de :

- Affiner la connaissance de la structure des populations au sein de la lignée Oryza à l'échelle mondiale grâce à une haute résolution de marqueurs et une forte couverture géographique ;
- Comprendre l'histoire de la propagation de l'agent pathogène à l'échelle globale ;
- Identifier les barrières au flux de gènes impliquées dans le maintien de la structure génétique observées à savoir :
    - l'allopatrie des lignées ;
    - les barrières à la reproduction sexuée incluant la compatibilité des types sexuels, la fertilité-femelle et les barrières post-appariement ;
    - l'adaptation différentielle des lignées à des facteurs environnementaux tels que l'hôte ou la température.

BARRIERS TO GENE FLOW BETWEEN PANDEMIC LINEAGES OF THE RICE BLAST FUNGUS *PYRICULARIA ORYZAE*

AUTHORS

Maud THIERRY, Joëlle MILAZZO, Henri ADREIT, Sonia BORRON, Violaine SELLA, Renaud IOOS, Elisabeth FOURNIER, Didier THARREAU, Pierre GLADIEUX

ABSTRACT

Natural variation in plant pathogens impacts food security and ecosystem health. The rice blast fungus *Pyricularia oryzae* is structured into multiple lineages and represents a major limitation to rice production in all rice-growing areas. Hypotheses for the origin of diversification and maintenance of multiple rice blast lineages include separation in different areas and differential adaptation to rice subspecies. However, the precise world distribution of rice blast populations, and factors controlling their presence and maintenance in the same geographic areas remain largely unknown. Here, we show that *P. oryzae* populations are structured into pandemic lineages, genetically isolated by breeding system, adaptation to hosts and separation in distinct climatic areas. We genotyped 886 isolates representing more than 185 locations in 51 countries using a SNP-genotyping chip and we found four pandemic lineages distributed in areas where distinct environmental conditions prevail, and different types of rice are grown. Differences in host range among lineages indicated that adaptation to local conditions contributes to genetic isolation between lineages. Furthermore, we found that ecology-independent premating or early postmating barriers, including female sterility, prevent successful mating between lineages. Our results demonstrate that the spread of a pathogen across heterogeneous habitats and divergent populations of a crop species can lead to incipient speciation in the pathogen. The design of new strategies for rice blast control using resistant varieties will need to take account of the strong genetic structure in *P. oryzae*.

INTRODUCTION

Understanding and controlling natural variation in fungal plant pathogens is of great importance for food security and ecosystem health. In addition to the overwhelming phylogenetic diversity of pathogens (Burdon, 1993), many pathogens also harbor substantial diversity at the species scale (Taylor and Fisher, 2003). Most fungal plant pathogens are subdivided into distinct populations (Taylor et al., 2006), and such population structure is a complex outcome of migration, selection and drift acting at multiple scales, and patterned by factors such as time (Ali et al., 2014; Pagliaccia et al., 2018), admixture (Robert et al., 2015), geography (Ali et al., 2014; Amend et al., 2009; Marin et al., 2009; Sork and Werth, 2014; Yang et al., 2018), reproductive mode (Ali et al., 2014; Büker et al., 2013; Gibson et al., 2012; Ropars et al., 2016) or environmental conditions (Walker et al., 2015). Among the environmental factors structuring populations of fungal plant pathogens, the identity of the host plants they interact with is often thought to have the strongest impact. This is because many plant pathogens reproduce within or on their host, which makes mating assortative with respect to host use and induces a strong association between adaptation to host and reproductive isolation (Giraud et al., 2010; Servedio et al., 2011). However, more generally, population structure can result from limited dispersal (i.e. limited gene flow because of distance) or limited adaptation (i.e. limited gene flow because of differences in the capacity to exploit resources), and both can be underlain by a wealth of potential factors that have remained largely unexplored in plant pathogens. Knowledge of how such barriers to gene flow contribute to the differentiation of populations within fungal pathogen species is a major goal for evolutionary microbiology because it informs our understanding of the origin of fungal biodiversity. While many studies have reported on the structure of fungal plant pathogens, none has done so based on a high density of samples and genetic markers, whereas barriers to gene flow have rarely been measured between populations.

Pyricularia oryzae (Ascomycota) is a model of widespread fungal plant pathogen with population subdivision. Several host-specific lineages, each mostly associated with one main cereal/grass host, have been characterized within *P. oryzae* (Gladieux et al. 2018a). Among these host-associated lineages, the lineage infecting Asian rice (*Oryza sativa*) has been extensively characterized. Previous studies of population structure in the rice-infecting lineage showed population subdivision, with the number of lineages ranging from three (Saleh et al., 2014; Tharreau et al., 2009; Zhong et al., 2018) to six (Gladieux et al. 2018b), and varying geographic patterns across lineages. The lineages were estimated to have diverged approximately 1000 years ago (Gladieux et al. 2018b; Zhong et al. 2018), and pathogenicity tests suggested local adaptation of some lineages to indica or japonica subspecies of rice (Gallet et al., 2016; Gladieux et al., 2018b). *P. oryzae* is a heterothallic fungus whose sexual cycle

involves mating between individuals of opposite mating-types, with at least one of the partners capable of producing female structure (i.e. “female-fertile”). Only one lineage, the lineage1 prevailing in Southeast Asia, shows a genome wide signal of recombination, balanced ratio of mating-type alleles and high frequency of fertile females, consistent with sexual reproduction (Saleh et al. 2012; Gladieux et al. 2018b). Other lineages displayed clonal population structure and frequencies of mating types or fertile females suggesting strictly asexual reproduction (Gladieux et al. 2018b; Saleh et al. 2014). Although previous studies have contributed to make *P. oryzae* a model for studying the population biology of fungal pathogens, most efforts to understand the population structure of the pathogen have been unable to provide an overview of the distribution of rice-infecting lineages at large scale, and the underlying phenotypic differences, either because sets of genetic markers were limited (Saleh et al., 2014; Tharreau et al., 2009), number of isolates was relatively small (Gladieux, et al. 2018b; Zhong et al. 2018) or because few phenotypic traits were scored (Saleh et al., 2014; Tharreau et al., 2009; Zhong et al., 2018).

Here we report on a detailed genomic and phenotypic overview of the diversity of rice-infecting *P. oryzae* pathogens sampled across all rice-growing areas of the world. Our main objective was to infer the population structure of the pathogen with high resolution in terms of markers and geographical coverage, to gain insight into the colonization history of the pathogen and to disentangle the barriers to gene flow underlying population structure. The specific questions addressed were the following: (i) what is the number of distinct lineages of *P. oryzae* infecting rice? (ii) What is the geographic distribution of rice-infecting lineages, and can we confirm previous findings about population structure? (iii) What factors could act as barriers to gene flow between lineages and thus explain the observed population structure? To address these questions, we used an Illumina beadchip microarray to score 5,657 SNPs distributed throughout the genome of 886 *P. oryzae* isolates collected on cultivated Asian rice in 51 countries. We also assessed barriers to gene flow caused by allopatry, female sterility, genetic incompatibilities and measured phenotypic divergence between lineages in terms of adaptation to hosts and temperature.

RESULTS

Rice-infecting *P. oryzae* populations are subdivided into four lineages

To characterize the global genetic structure of the rice blast pathogen, 886 *P. oryzae* isolates were genotyped for 5,657 SNPs markers distributed throughout the genome, resulting in 3,686 SNPs after filtering out positions with low quality or missing data. Clustering analyses based on sparse nonnegative matrix factorization algorithms, as implemented in the program sNMF, uncovered four

clusters, hereafter referred to as "lineages" (Figure 1 a, b). We selected the model with K=4 clusters as the best model based on the cross-entropy criterion, as models with K>4 only induced a small decrease in cross-entropy, suggesting that K=4 captures the most salient features of the structure of the dataset (Supplementary figure 1). Neighbor phylogenetic network inferred with Splitstree also supported subdivision into four lineages, with long branches separating three peripheral lineages that are branching inside a central lineage (Figure 1 a). Comparison with previous findings reveals that the three peripheral lineages in the network corresponded to previously described lineages 2, 3 and 4 (Gladieux et al. 2018b), while the central lineage corresponded to the combination of recombining lineage 1 and the lineages 5 and 6 represented by few individuals in Gladieux et al. (2018b). The central lineage is subsequently referred to as lineage 1, for the sake of simplicity. Lineages 2 and 3 were similar to the lineages B and C, respectively, previously identified using microsatellite data (Saleh et al., 2014 - Supplementary figure 1).

Genetic differentiation between the four lineages was high and significant (Weir and Cockerham's F_{ST}> 0.54), indicating strong barriers to gene flow between them. While all genotypes from lineages 2 to 4 had high membership proportions *q* in a single cluster (*q* > 0.89), admixture was detected in lineage 1 with the three other lineages, with 32% of genotypes having *q* > 0.10 in lineages 2-4. Admixture may account for the lower F_{ST} observed in comparisons between lineage 1 and other lineages (Figure 1).

Admixture and geographical substructure in recombining lineage 1

Lineage 1 was the only lineage to display population genetic evidence for recombination (Phi-test testing the null hypothesis of clonality: p-value =0.00). Most isolates from lineage 1 were collected in Asia (79%) but the lineage was also present wherever the pathogen was sampled (Europe: 1 isolate; North America: 10, Central and South America: 7, Africa: 22) (Figure 2). Clustering analysis of this single lineage with sNMF detected four sub-clusters with different geographic distributions (Figure 2). Estimates of differentiation were lower between the sub-clusters of lineage 1 (F_{ST} < 0.54) than between the main lineages (F_{ST} > 0.49) consistent with a longer history of restricted gene flow between main lineages (Supplementary table 1). Two sub-clusters (hereafter referred to as 1-bao and 1-yule) were mostly composed of isolates sampled in two distinct sites from Yunnan (respectively Baoshan and Yule) distant of c.a. 700 km. The third sub-cluster was mostly composed of isolates from Laos and South China (hereafter referred to as 1-laos), while the fourth sub-cluster gathered 95% of the isolates from lineage 1 collected outside Asia (hereafter referred to as 1-int). In Asia, sub-cluster 1-int was mostly present in the Yunnan province of China, India and Nepal (Figure 2). Admixture was widespread in

lineage 1 (Figure 2) and most of the isolates (78%) displayed membership proportions $q > 0.10$ in two or more sub-clusters (Figure 2). The only five genotypes detected in multiple countries among all genotypes (genotype ID [number of countries]: 2 [6], 18 [2], 58 [2], 98 [3] and 254 [3]; representing 41 isolates in total) all belonged to sub-cluster 1-int.

Colonization history of clonal lineages 2, 3 and 4

No population genetic signals of recombination were detected in lineages 2, 3 and 4 (phi-test p-values: 0.25, 0.18 and 0.22, respectively), confirming previous findings (Saleh et al. 2014; Gladieux et al. 2018b; Zhong et al. 2018). The lack of recombination enabled the use of a phylogenetic approach using the RAxML program to infer the relationships among genotypes within lineages (Figure 3). The earliest branching isolates of all three lineages were Asian, which is consistent with an Asian origin of clonal lineages (Figure 3). All genotypes from Madagascar formed a clade nested within lineage 3, concordant with a single introduction of this lineage in the island. European genotypes were scattered across the genealogy of lineage 2 with extensive sharing of genotypes among countries suggesting multiple introductions and intense movements of biological material from, to, or within Europe.

Shared genotypes suggest long-distance movement of the pathogen

In all lineages, some isolates from different countries, including isolates from different continents, shared identical genotypes (i.e. displayed the same allelic profile at 3,686 SNPs). Although a minority of the shared genotypes may result from convergence and homoplasy, most shared genotypes likely result from the direct or indirect introductions of genotypes, probably through exchange of contaminated material between countries. Analyzing the number of shared genotypes between countries revealed the historical relationships among regions in terms of movements of *P. oryzae* (Figure 4). Plotting shared genotypes onto a map suggested extensive movement of the pathogen within Europe, America or Asia, but not within Africa with our sampling. At the intercontinental level, only Brazil and the Philippines shared three or more genotypes (one genotype from lineage 1, one genotype from lineage 2 and one genotype from lineage 3), suggesting direct or indirect exchange of strains between the two countries.

Pre-mating barriers caused by allopatry

We described four major lineages (lineages 1 to 4) in the rice blast pathogen with high differentiation between them (F_{st}>0.54) and admixture restricted to recombining lineage 1. We therefore sought to assess several possible barriers to gene flow that could be involved in the observed genetic structure. We first investigated allopatry, by analysing the geographic distribution of lineages.

Lineages 2 and 3 are both widespread but with substantial differences in their geographical range. Lineage 2 is present in all continents but is the only one sampled in Europe (except one isolate from lineage 1). Lineage 3 is present in all continents except Europe and is strongly represented in inter-tropical regions. Lineage 4 groups a smaller number of isolates but is present in three continents. Most genotypes from lineage 4 (70%) were found in India, and the remaining in other parts of Asia (Bangladesh, China, Nepal), in Africa (Benin, Tanzania) or in the USA. Finally, lineage 1 is predominant in Asia (79% of all isolates and 94% of all genotypes). Outside Asia, lineage 1 is present in all continents (Supplementary table 2). The country of origin of each isolate in this study is known, and a more precise sampling location is available for most (region, city or GPS coordinates). On several instances, two or even three distinct lineages were sampled the same year in the same city (pattern observed in 11 countries spread over all continents), showing that coexistence of the lineages in the same very restricted geographical area is not uncommon. However, the coexistence of two lineages at the exact same GPS position was rarely observed (with the exception of isolates US0106 and US0107 sampled from the same field). Allopatry of the lineages on a smaller scale is therefore not excluded.

We investigated if differences in geographical range of the lineages could be associated with differences in climatic variables. Plotting sampling locations onto the map of major climate regions (Kottek et al. 2006) indicated differences in the climatic distribution of lineages 2 and lineage 3, with lineage 2 mostly observed in warm temperate climates and lineage 3 in equatorial climates. To further test the hypothesis that lineages are distributed in regions with distinct climatic characteristics, we retrieved 19 climatic variables (named biomes) from the WorldClim bioclimatic variables database (Fick et Hijmans 2017) for all sampling locations (only if the region, city or GPS position of the sampling was known). We used the Outlying Mean Index (OMI), which measures the distance between the mean habitat condition used by a lineage and the mean habitat conditions used by the entire species. A statistical permutation test on OMI values demonstrated a significance of marginality for the lineages 2, 3 and 4 considering the environmental variable under study (p-values L1: 0.6213; L2: 0.0003; L3: 0.0001; L4: 0.0176). The first two axes of the OMI analysis explained respectively 69% and 25% of the variability. This analysis also highlighted the most important ecological factors differentiating the distribution of lineages. Lineage 2 was highly represented in regions with high temperature annual range (biome7) or high seasonality (biome4), lineage 4 was present in regions with high seasonal precipitations (biomes 13, 16 or 18), and lineage 3 was overrepresented in regions with high temperature (biomes 1, 6, 10 and 11) and high isothermality (biome3) characteristic of tropical climates (Figure 5 b and c; Supplementary figure 3).

Ecologically-independent pre- and early post-mating reproductive barriers caused by female sterility and genetic incompatibilities

We evaluated the capacity of the different lineages to engage into sexual reproduction, by characterizing mating type, female fertility and interfertility in vitro between isolates of opposite mating type. Lineages 2, 3 and 4 were almost single mating type (97% of lineages 2 and 4 isolates tested carry the *Mat1-1* allele; 97% of lineage 3 isolates carry the *Mat1-2* allele; remaining 3% are opposite mating types), suggesting asexual reproduction. The mating type ratio was more balanced in lineage 1 (52% of Mat1-1; Figure 6 A). These results suggest that for lineages 2-4 reproduction is potentially possible between lineages but not within lineages, whereas lineage 1 is potentially compatible with all lineages. However, the mating type is not the only physiological determinant of sexual compatibility, which also depends on the ability to produce female structures. The proportion of isolates with female fertility was low in lineages 2-4 (0-7% of female-fertile isolates), which reduces the opportunities for crossing between these lineages, despite their compatible mating types. In contrast, the proportion of isolates capable of producing female structures was much higher in some lineage 1 sub-clusters (1-yule: 79%; 1-bao: 67%; 1-laos: 37%), which is a favorable condition for outcrossing between lineages.

Compatibility within and between lineages was measured by scoring the formation of sexual structures (perithecia) in crosses between isolates of opposite mating-types randomly selected (Figure 6 B). This experiment revealed a marked heterogeneity in the rate of perithecia formation across lineages. Isolates in the 1-yule sub-cluster produced perithecia in 93% of crosses with isolates from the same sub-cluster, and in more than 46% of crosses with isolates from other lineages. Isolates from lineage 2, 3 and 4 could only be crossed with isolates from other lineages, given that isolates tested were single mating type, and the percentage of these crosses producing perithecia was highly variable (ranging from 0% to 83% depending on the lineages involved). The rate of perithecia formation was comparable in the 1-int sub-cluster of lineage 1 to what was observed in the clonal lineages 2-4 and none of the intra-sub-cluster 1-int crosses lead to perithecia formation. Other sub-clusters of lineage 1 (i.e. 1-bao and 1-laos) displayed intermediate rates of perithecia formation (Figure 6 B).

To further characterize reproductive compatibility within and between lineages, we opened perithecia and scored the formation of asci (i.e. meiotic octads) and the germination of ascospores (i.e. meiospores) for a subset of crosses involving some of the most fertile isolates. This experiment revealed that most inter-lineage crosses produced perithecia that did not contain asci or containing asci with reduced ascospores germination rate. While 100% of the crosses between strains of sub-cluster 1-yule produced numerous germinating ascospores, this rate was reduced to 33%, 56% and 7% in 1-yule x lineage 2, 1-yule x lineage 3 and 1-yule x lineage 4, respectively. Together, the results presented in this

section indicate that the lineages in *P. oryzae* are isolated by strong ecologically-independent pre- and early-postmating barriers, including breeding system isolation (differences in mating type and female sterility), and a combination of gametic incompatibility, zygotic mortality or hybrid unviability.

Ecologically-dependent pre-mating and post-mating barriers caused by adaptation to hosts

We challenged 45 rice varieties representing the five main genetic groups of Asian rice, *Oryza sativa* (temperate japonica rice, tropical japonica, aus, aromatic and indica) with 70 isolates representing the four lineages of *P. oryzae* and the four sub-clusters within lineage 1 to test the hypothesis of adaptation to the host (Figure 7). Qualitative symptoms notations were analysed using a proportional-odds model that revealed significant differences between lineages (p-value = $2.2x10^{-16}$), between rice genetic groups (p-value=$2.2x10^{-16}$) as well as a significant interaction between these variables (p-value=$2.0x10^{-10}$). In comparisons between lineages, lineage 2 stood out from the other lineages, as isolates from this lineage generally produced smaller lesions (Figure 8 a). Considering rice genetic groups, significantly strongest symptoms were observed on temperate japonicas, while the varieties of the aromatic genetic group were significantly more resistant to rice blast (Figure 8 b). The finding of a significant interaction between lineages and rice genetic groups indicated that the effect of the lineage of origin of isolates on the proportion of compatible interactions differed between the rice types and suggests adaptation to hosts. Analyses of interactions also revealed high pathogenicity of the 1-yule sub-cluster toward most indica varieties (Figure 7). All isolates from this sub-cluster were pathogenic to all indica varieties but IR8. The IR8 variety was one of the most resistant varieties in our panel, with only 15 isolates pathogenic to this variety (i.e. symptoms notation >2), 47% of which belong to lineage 3. Together, these experiments reveal significant differences in host ranges between lineages, but all host ranges were overlapping, indicating that specialization to the host is not strict. Adaptation to the host could nonetheless contribute to further reduce gene flow between populations, via pre-mating barriers (unviability of immigrants, i.e. reduced encounters between potential mates due to mortality of immigrants - Gladieux et al. 2011; Rundle et Nosil 2005) and post-mating barriers (ecological hybrid unviability, i.e. reduced survival of ill-adapted hybrid offspring).

Ecologically-dependent pre-mating and post-mating barriers caused by adaptation to temperature

We measured growth rate and sporulation rate of representative isolates cultured at different temperatures to test the hypothesis of adaptation to temperature. For all lineages, mycelium growth rate increased with incubation temperature, although this trend was more obvious from 10°C to 15°C (increased mean mycelium growth of +2.22 mm/day) than from 25°C to 30°C (+0.05 mm/day) (Figure 9). Comparing lineages at each temperature evidenced a significantly lower growth rate of lineage 4 at

10°C than other lineages (p-value< 0.02) and a significantly higher growth rate of lineage 1 (sub-cluster 1-yule) at 15°C, 20°C, 25°C and 30°C compared to other lineages (p-value< 0.05) (Figure 9). Likewise, the in vitro sporulation rate was significantly affected by temperature (Figure 10). Sporulation was nearly absent at 10°C after 20 days of culture. The few spores observed were often not completely formed and were only divided by one septum instead of two septa in mature conidia. Regardless of the lineage, sporulation rate increased with temperature from 15 to 25 °C. At 30°C sporulation rates dropped for all lineages but isolates were cultured 7 days only, instead of 10 days at 25°C. At 30°C, a significant effect of the lineages was observed (p-value = 0.036). However, pairwise comparison of lineages did not reveal significant differences. At 15°C, sporulation was significantly higher for lineages 2, 3 and 4 compared to lineage 1 (p-value <0.0001).

In summary, our results suggest that there is no strong measurable barrier to gene flow due to the effect of temperature on immigrant viability or ecological hybrid viability (because growth and sporulation showed comparable trend for all lineages at all temperatures), and therefore that adaptation to temperature may not play a major role in the genetic isolation of lineages. However, the observed differences in sporulation and growth rate suggest that some lineages might outcompete others at certain temperatures, and thus reduce the likelihood of encounters.

DISCUSSION

We report on the existence of strong barriers to gene flow between lineages of *P. oryzae* causing rice blast. In our analysis of SNP-genotyping data, we have shown that *P. oryzae* is subdivided into four lineages. Two of the lineages previously detected using whole genome data (lineages 5 and 6; Gladieux et al., 2018b), and represented by few individuals included in our dataset, were assigned to lineage 1 in our analysis, and could correspond in our dataset to the subdivision of lineage 1. Differences in the number of lineages detected could be due to an ascertainment bias, given that our SNP-genotyping beadchip was designed using a set of genomic sequences in which representatives of lineages 5 and 6 had not been included. Another hypothesis could be that the sampling used in the study of Gladieux et al. (2018b) was not dense enough to highlight the continuum existing between lineages 1, 5 and 6, making them appear as discrete entities. Consistent with previous findings, analyses of phylogenetic relationship among isolates revealed three clonal clusters connected by long branches to a central recombining cluster. Some genotypes assigned to lineage 1 might eventually in the future resemble other clonal lineages in phylogenetic analyses, and form a long branch stemming from central lineage 1, especially genotypes already propagated clonally such as genotype 2, 18, 58, 98 or 254. Only one of the 24 non-Asian isolates of lineage 1 for which fertility was measured was able to produce

female structures, indicating that propagation out of Asia is strongly associated with a loss of fertility in *P. oryzae*.

We evaluated several potential barriers to gene flow between lineages 1 to 4. Female sterility, mating-type incompatibilities and ecologically-independent pre- and post-mating barriers (i.e., intrinsic incompatibilities) appear to be strong barriers to gene flow between clonal lineages 2, 3 and 4. However, some of the attempted crosses between clonal lineage 2-4 and recombining lineage 1 produced viable progeny, confirming the possibility of gene flow into this lineage, as previously evidenced based on admixture mapping (Gladieux et al., 2018b). Our analyses also revealed allopatry and separation into distinct climatic areas as another potentially strong barrier to gene flow between lineages. Lineage 1 is mostly sampled in Southeast Asia, lineage 4 is predominant in India, while lineages 2 and 3 are pandemic. Analysis of climatic data indicated that lineage 2 is predominant in temperate climates, where temperate japonica rice is grown, lineage 3 in tropical climates, where indica rice is grown, and lineage 4 in regions with high seasonal precipitations where indica and aromatic rice types dominate. Despite the finding of separation in different climatic regions, our experiments revealed no strong differences between lineages in terms of sporulation and mycelial growth on synthetic media at different temperatures. This suggests that if adaptation to temperature exists in this pathogen, it is not measured by our experiments, either because other traits than sporulation and hyphal growth are involved or in vitro conditions were not suitable to evidence differences. The host range varied across lineages but all host ranges were overlapping, indicating that specialization to the host is not strict. Comparable overlapping host ranges were previously highlighted by Gallet et al. (2016). Adaptation to the host could nonetheless contribute to further reduce gene flow between populations. For instance, lineage 2 displayed a narrow host range which may reduced opportunities for encounters and mating with other groups. Finally, we note that eleven out of the 12 most multivirulent isolates (symptoms > 2 on more than 40 varieties tested) belong to lineage 1. The propagation of such genotypes, which belong to a lineage that has both mating types and fertile females, should be monitored closely.

MATERIAL AND METHODS

Biological material

Eight hundred eighty-six *P. oryzae* isolates collected on Asian rice between 1954 and 2014 were selected to represent the global genetic diversity of the fungus. Isolates were selected based on microsatellite data, in order to maximize the number of multilocus genotypes represented (Adreit et al. 2007; Saleh et al. 2014; Odjo et al. in press, and unpublished data), or based on their geographic

origin when genotypic data were not available, in order to maximize the number of countries represented in the dataset.

Sixty-eight isolates were selected to carry out experimental measures of reproductive success, adaptation to host, growth and sporulation experiments at different temperatures (Supplementary table 2). This subset of isolates included 10 isolates from each of the three clonal lineages (lineages 2-4), 27 isolates from the different sub-clusters within lineage [1-bao (9 isolates), 1-int (10), 1-laos (8), and 1-yule (10)] and isolate CH0999, which is a reference female-fertile strain with Mat1-1 mating-type (Saleh et al., 2012a).

Forty six varieties were chosen as representative of the 5 main gene pools of Asian rice (Garris et al., 2005): indica (Chau, Chiem chanh, DA11, De abril, IR8, JC92, JC120, Pappaku), aus (Arc 10177, Baran boro, Black gora, DA8, Dholi boro, Dular, FR13 A, JC148, Jhona 26, Jhona 149, Kalamkati, T1, Tchampa, Tepi boro), temperate japonica (Aichi asahi, Kaw luyoeng, Leung pratew, Maratelli, Nep hoa vang, Nipponbare, Sariceltik, Som Cau 70A), tropical japonica (Azucena, Binulawan, Canella de ferro, Dholi boro, Gogo lempuk, Gotak gatik, Moroberekan, Trembese) and aromatic (Arc 10497, Basmati lamo, Dom zard, Firooz, JC1, Kaukkyisaw, N12). Varieties Maratelli (temperate japonica) and CO39 (indica) were used as susceptible controls.

Genotyping

P. oryzae isolates were genotyped at 5,657 genomic positions by Bayer Crop Science Singapore using an Illumina Infinium beadchip microarray designed to focus on single nucleotide polymorphisms (SNPs) identified in 25 *P. oryzae* genomes previously sequenced (Gladieux et al., 2018b). According to our clustering analysis, the 25 isolates belong to lineage 1 (n=2 isolates), lineage 2 (n=9), lineage 3 (n=8) or lineage 4 (n=3), and seven isolates were not characterized here. The final dataset included 3,686 SNPs biallelic with no missing data.

Population subdivision and recombination

Among the 886 *P. oryzae* isolates genotyped, we identified 264 distinct multilocus genotypes that were used for the analysis of population subdivision. We used the program sNMF to infer individual ancestry coefficients in K ancestral populations. This program is optimized for large dataset analysis and avoids Hardy-Weinberg equilibrium assumptions, which makes it more appropriate to deal with inbred or clonal lineages (Frichot et al., 2014). We used the SPLITSTREE program to visualize relationships among genotypes in a phylogenetic network, with reticulations representing conflicting

phylogenetic signals caused by homoplasy. We also used the Pairwise Homoplasy Index (PHI) test implemented in SPLITSTREE to test the null hypothesis of clonality. We then inferred a maximum likelihood genealogy to further investigate the relationships among genotypes within the clusters for which clonality could not be rejected, and we used the sNMF software to further investigate population subdivision within the cluster for which the null hypothesis of clonality could be rejected. Maximum likelihood genealogies were inferred using RAxML 8.2.4 with the GTRGAMMA mode and 100 bootstraps values. Weir and Cockerham's F_{ST} was calculated using the hierfstat packages implemented in R software with the WC84 method (Goudet, 2005).

Shared MLG representation

We used global geographic data "world.gpkg" from the spData package implemented in the sf library (version 0.7-4) in R. The cartography library (version 2.2.0 – Giraud and Lambert 2017) was used to represent shared MLGs between countries.

Experimental measures of reproductive success and female fertility

P. oryzae is a heterothallic fungus with 2 mating-types (Mat1-1 and Mat1-2). Sexual reproduction between strains of opposite mating type can be observed in lab conditions and results in the production of ascospores inside female sexual structures called perithecia (Saleh et al. 2012). On synthetic media, perithecia are formed at the contact zone between parental mycelia. Crosses were carried out on a rice flour agar medium (20 g rice flour, 2 g yeast extract, 15 g agar and 1 L water, supplemented with 500 000 IU of penicillin G after autoclaving for 20 min at 120°C) as described by Saleh et al. (2012). We assessed reproductive success by measuring the amount of perithecia produced after 3 weeks of cultivation at 20°C under continuous light (0: no perithecium formed; 1: less than 10 perithecia; 2: more than 10 perithecia at some parts of the contact zone between the 2 mycelia; 3: perithecia all along the contact zone). Perithecia can be formed by the two interacting partners or by only one. Isolates forming perithecia are qualified as female fertile. We measured female fertility of some strains (listed in Supplementary table 2) by monitoring perithecia production in crosses involving tester strains CH0997 (Mat1-2) and CH0999 (Mat1-1). We further assessed the presence of asci and germinating ascospores in perithecia for a subset of crosses by excising perithecia with a scalpel and counting the number of germinated filament for each individualized ascus after overnight incubation on water agar. The subset of crosses included 10 Mat1-1 isolates (lineage 1: CH0999, CH1065, CH1076; lineage 2: CH0092, MC0016, SP0006; lineage 4: IN0017, IN0092, NP0070, CH0718) and 6 Mat1-2 isolates (lineage 1: CH0997; CH1083, CH1120; lineage 3: BR0019, CH0549, MD0929).

Pathogenicity tests

Compatibility between *P. oryzae* isolates and rice plants representing the five main genetic groups of rice (indica, temperate japonica, tropical japonica, aus and aromatic) was measured in controlled conditions. Seventy isolates were inoculated on 46 varieties. Inoculations were performed as described by Gallet et al. (2016). Conidial suspensions (25 000 conidia.mL-1) with 0.5% gelatin were sprayed on 3 weeks old rice seedlings (> 6 plants/variety). The inoculated plants were incubated 8 hours at 27°C and 100% humidity then 7 days with a day/night alternation (13 hours at 27°C / 11 hours at 21°C) before scoring symptoms. Inoculations and scoring of symptoms were repeated three times. Lesion type was rated from 1 to 6 (Gallet et al. 2016) and the disease area was assessed visually on leaves.

Mycelial growth and sporulation rate at different temperatures

Mycelial growth and sporulation rate was measured for 42 isolates at 5 different temperatures (10°C, 15°C, 20°C, 25°C and 30°C). For each isolate, Petri dishes with PDA medium were inoculated with mycelial plugs placed at the center and incubated in a cultivation chamber with a fixed temperature. Mycelium diameter was measured along 2 perpendicular axes at different time points. At the end of the experiment, conidia were collected by adding 5 mL of water with 0.01 % of tween 20 in the Petri dish and rubbing the surface of the mycelium. Conidia were counted using Malassez cells. Three repetitions were carried out for each isolate at each temperature.

Statistical analyses of climatic and phenotypic data

Ecological niche separation: OMI (Outlying Mean Index), or marginality, is used to study niche separation and niche breadth. It gives the same weight to all samplings whether rich or poor in species and individuals and is particularly suitable in cases where the sampling is not homogeneous. This index measures the deviation between the mean environmental conditions used by one lineage and the mean environmental conditions used by all lineages. OMI analysis then places the lineages along environmental conditions maximizing their OMI. To each sampling location were associated environmental values (WorldClim bioclimatic variables – Fick and Hijmans 2017) consisting of 19 biomes values. Climatic values were normalized before the analysis. A contingency table associated the number of isolates from each lineage to each sampling location. Only isolates with a precise location (known region, city or GPS position of the sampling) were included. A random permutation test, with 10000 permutations, checks the statistical significance of the marginality for each lineage.

Pathotyping: Ordinal symptoms notation was analysed using a proportional-odds model which accounts for ordered categorical responses using clm() function implemented in the ordinal R package

(version 2018.8-25). Significance of the factors was analyzed using an Anova. Pairwise comparisons of significant factors were done after computing least-squares means using lsmeans (lsmeans R package version 2.30-0) and using a Tukey adjustment.

Mycelium growth: Mycelium growth was analyzed for each temperature independently with a linear mixed-effect model. Significance of the factors was analyzed using an ANOVA. Pairwise comparisons of significant factors were done after computing least-squares means using lsmeans (lsmeans R package version 2.30-0).

Sporulation: Median value of the number of spores calculated for the different repetitions were analyzed using a negative binomial generalized linear model. Pairwise comparisons of significant factors were done by computing least-squares means.

LITTERATURE CITED

Adreit, H., Andriantsimialona, D., Utami, D.W., Notteghem, J.L., Lebrun, M.H., and Tharreau, D. (2007). Microsatellite markers for population studies of the rice blast fungus, *Magnaporthe grisea*. Mol. Ecol. Notes *7*, 667–670.

Ali, S., Gladieux, P., Rahman, H., Saqib, M.S., Fiaz, M., Ahmad, H., Leconte, M., Gautier, A., Justesen, A.F., Hovmøller, M.S., et al. (2014). Inferring the contribution of sexual reproduction, migration and off-season survival to the temporal maintenance of microbial populations: a case study on the wheat fungal pathogen *Puccinia striiformis* f.sp. *tritici*. Mol. Ecol. *23*, 603–617.

Amend, A.S., Garbelotto, M., Fang, Z., and Keeley, S.C. (2009). Isolation by landscape in populations of a prized edible mushroom Tricholoma matsutake. Conserv. Genet. *11*, 795–802.

Büker, B., Petit, E., Begerow, D., and Hood, M.E. (2013). Experimental hybridization and backcrossing reveal forces of reproductive isolation in *Microbotryum*. BMC Evol. Biol. *13*, 224.

Burdon, J.J. (1993). The Structure of Pathogen Populations in Natural Plant Communities. Annu. Rev. Phytopathol. *31*, 305–323.

Fick, S.E., and Hijmans, R.J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. *37*, 4302–4315.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. (2014). Fast and Efficient Estimation of Individual Ancestry Coefficients. Genetics *196*, 973–983.

Gallet, R., Fontaine, C., Bonnot, F., Milazzo, J., Tertois, C., Adreit, H., Ravigné, V., Fournier, E., and Tharreau, D. (2016). Evolution of Compatibility Range in the Rice− *Magnaporthe oryzae* System: An Uneven Distribution of R Genes Between Rice Subspecies. Phytopathology *106*, 348–354.

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic Structure and Diversity in *Oryza sativa* L. Genetics *169*, 1631–1638.

Gibson, A.K., Hood, M.E., and Giraud, T. (2012). Sibling Competition Arena: Selfing and a Competition Arena Can Combine to Constitute a Barrier to Gene Flow in Sympatry. Evolution *66*, 1917–1930.

Giraud, T., and Lambert, N. (2017). Reproducible Cartography. In Advances in Cartography and GIScience, M.P. Peterson, ed. (Springer International Publishing), pp. 173–183.

Giraud, T., Gladieux, P., and Gavrilets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. Trends Ecol. Evol. *25*, 387–395.

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J.L.N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., et al. (2018a). Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the Rice Blast Fungus *Magnaporthe oryzae*. MBio *9*, e01219-17.

Gladieux, P., Ravel, S., Rieux, A., Cros-Arteil, S., Adreit, H., Milazzo, J., Thierry, M., Fournier, E., Terauchi, R., and Tharreau, D. (2018b). Coexistence of Multiple Endemic and Pandemic Lineages of the Rice Blast Pathogen. Mbio *9*, e01806-17.

Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. Mol. Ecol. Notes *5*, 184–186.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. Meteorol. Z. *15*, 259–263.

Marin, M., Preisig, O., Wingfield, B.D., Kirisits, T., and Wingfield, M.J. (2009). Single sequence repeat markers reflect diversity and geographic barriers in Eurasian populations of the conifer pathogen *Ceratocystis polonica*. For. Pathol. *39*, 249–265.

Pagliaccia, D., Urak, R.Z., Wong, F., Douhan, L.I., Greer, C.A., Vidalakis, G., and Douhan, G.W. (2018). Genetic Structure of the Rice Blast Pathogen (Magnaporthe oryzae) over a Decade in North Central California Rice Fields. Microb. Ecol. *75*, 310–317.

Robert, S., Zapater, M., Carlier, J., Abadie, C., and Ravigné, V. (2015). Multiple introductions and admixture at the origin of the continental spread of the fungal banana pathogen *Mycosphaerella fijiensis* in Central America: a statistical test using approximate bayesian computation. p.

Ropars, J., Lo, Y., Dumas, E., Snirc, A., Begerow, D., Rollnik, T., Lacoste, S., Dupont, J., Giraud, T., and López-Villavicencio, M. (2016). Fertility depression among cheese-making *Penicillium roqueforti* strains suggests degeneration during domestication. Evol. Int. J. Org. Evol. *70*, 2099–2109.

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigné, V., Bazin, E., Nottéghem, J.-L., Fournier, E., et al. (2012a). Sex at the origin: an Asian population of the rice blast fungus *Magnaporthe oryzae* reproduces sexually. Mol. Ecol. *21*, 1330–1344.

Saleh, D., Milazzo, J., Adreit, H., Tharreau, D., and Fournier, E. (2012b). Asexual reproduction induces a rapid and permanent loss of sexual reproduction capacity in the rice fungal pathogen *Magnaporthe oryzae*: results of in vitroexperimental evolution assays. BMC Evol. Biol. *12*, 42.

Saleh, D., Milazzo, J., Adreit, H., Fournier, E., and Tharreau, D. (2014). South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. New Phytol. *201*, 1440–1456.

Servedio, M.R., Van Doorn, G.S., Kopp, M., Frame, A.M., and Nosil, P. (2011). Magic traits in speciation: “magic” but not rare? Trends Ecol. Evol. *26*, 389–397.

Sork, V.L., and Werth, S. (2014). Phylogeography of Ramalina menziesii, a widely distributed lichen-forming fungus in western North America. Mol. Ecol. *23*, 2326–2339.

Taylor, J.W., and Fisher, M.C. (2003). Fungal multilocus sequence typing — it's not just for bacteria. Curr. Opin. Microbiol. *6*, 351–356.

Taylor John W, Turner Elizabeth, Townsend Jeffrey P, Dettman Jeremy R, and Jacobson David (2006). Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. Philos. Trans. R. Soc. B Biol. Sci. *361*, 1947–1963.

Tharreau, D., Fudal, I., Andriantsimialona, D., Santoso, Utami, D., Fournier, E., Lebrun, M.-H., and Notteghem, J.-L. (2009). World Population Structure and Migration of the Rice Blast Fungus, *Magnaporthe oryzae* (Dordrecht: Springer).

Walker, A.-S., Gladieux, P., Decognet, V., Fermaud, M., Confais, J., Roudet, J., Bardin, M., Bout, A., Nicot, P.C., Poncet, C., et al. (2015). Population structure and temporal maintenance of the multihost fungal pathogen *Botrytis cinerea*: causes and implications for disease management. Environ. Microbiol. *17*, 1261–1274.

Yang, N., Ma, G., Chen, K., and Wu, X. (2018). The Population Genetics of *Alternaria tenuissima* in Four Regions of China as Determined by Microsatellite Markers Obtained by Transcriptome Sequencing. Front. Microbiol. *9*.

Zhong, Z., Chen, M., Lin, L., Han, Y., Bao, J., Tang, W., Lin, L., Lin, Y., Somai, R., Lu, L., et al. (2018). Population genomic analysis of the rice blast fungus reveals specific events associated with expansion of three main clades. ISME J.

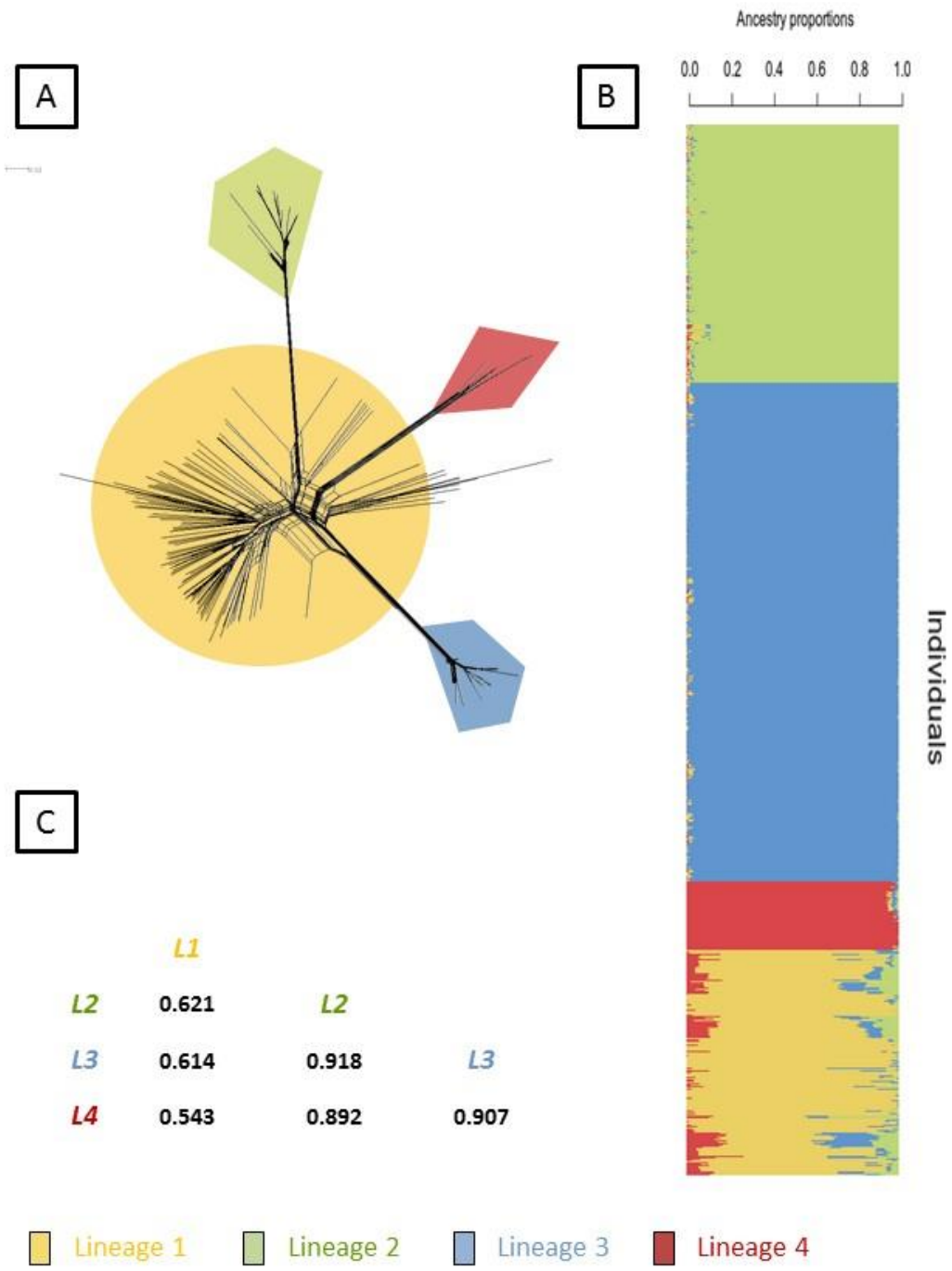


Figure 1: Population subdivision in 264 *P. oryzae* multilocus genotypes, representing 886 isolates, with four inferred lineages represented in different colors. A: Neighbor-net phylogenetic network estimated with Splitstree; reticulations indicate phylogenetic conflicts caused by homoplasy B: Ancestry proportions in K=4 clusters, as estimated using the sNMF software; C: Pairwise Weir and Cockerham FST between lineages.

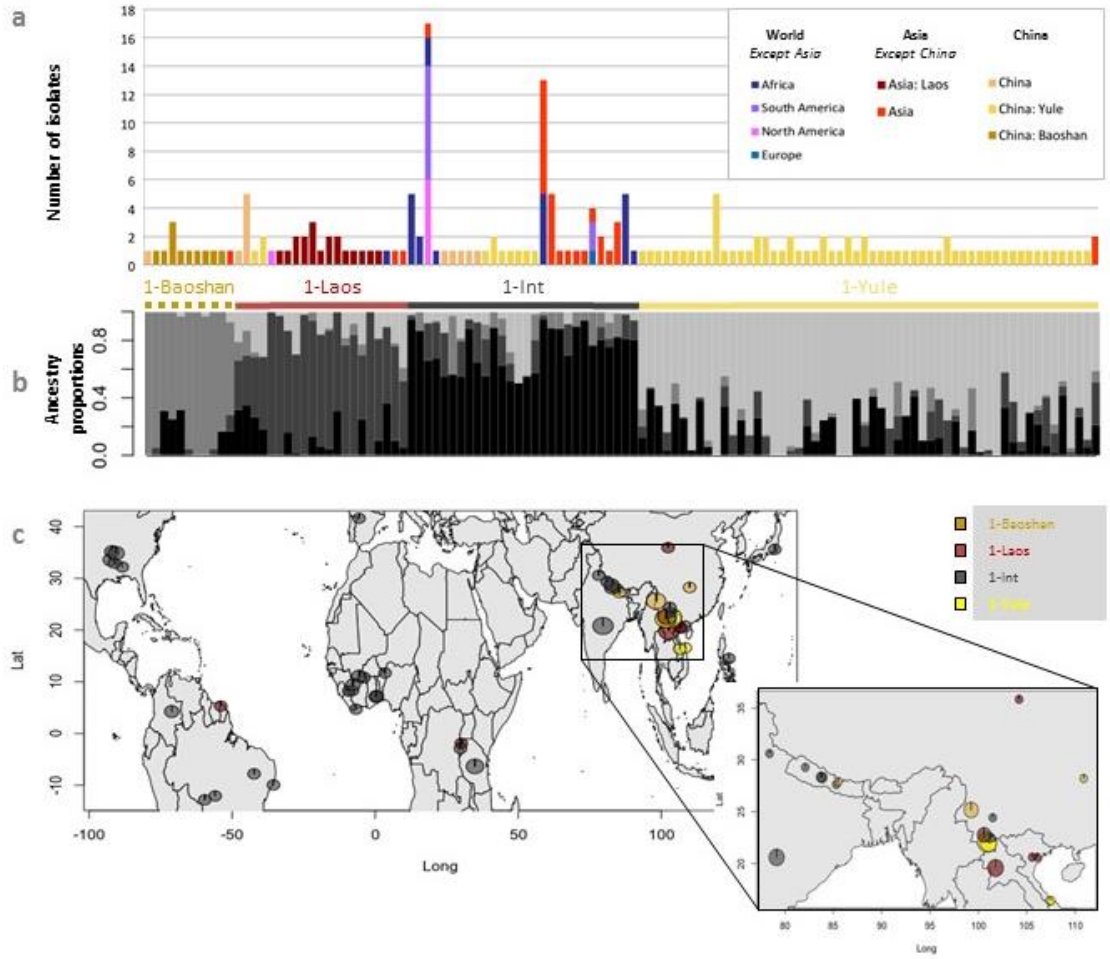


Figure 2: Population subdivision within lineage 1. a: Number of isolates and their geographic origin for each multilocus genotype of the lineage 1; b: Ancestry proportions in four sub-clusters with information about the geographic origin of isolates; c: World distribution of the four sub-clusters identified.

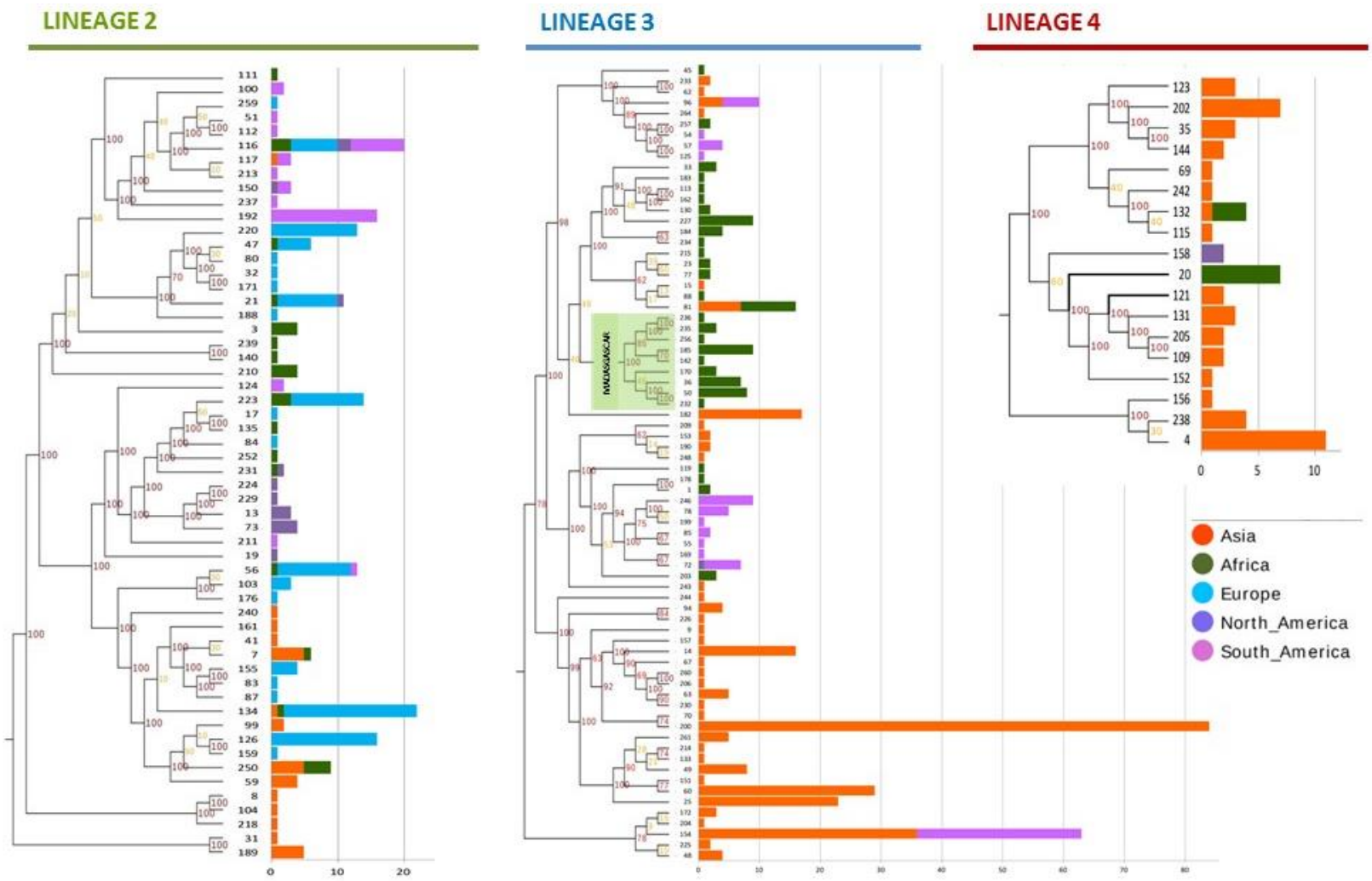


Figure 3: Maximum likelihood genealogies showing the relationships between genotypes from lineages 2 to 4 (RaXml 8.2.4). Barplots represent the number of isolates representing each genotype and colors represent their continent of origin. The green square highlights the branch in lineage 3's genealogy grouping all Madagascan isolates.

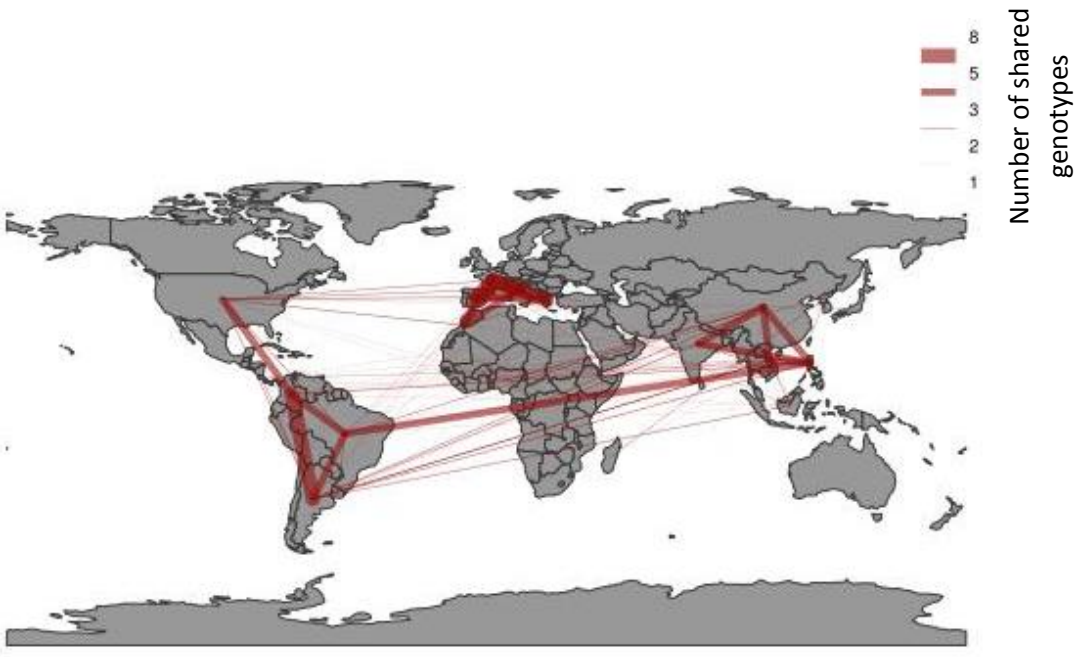
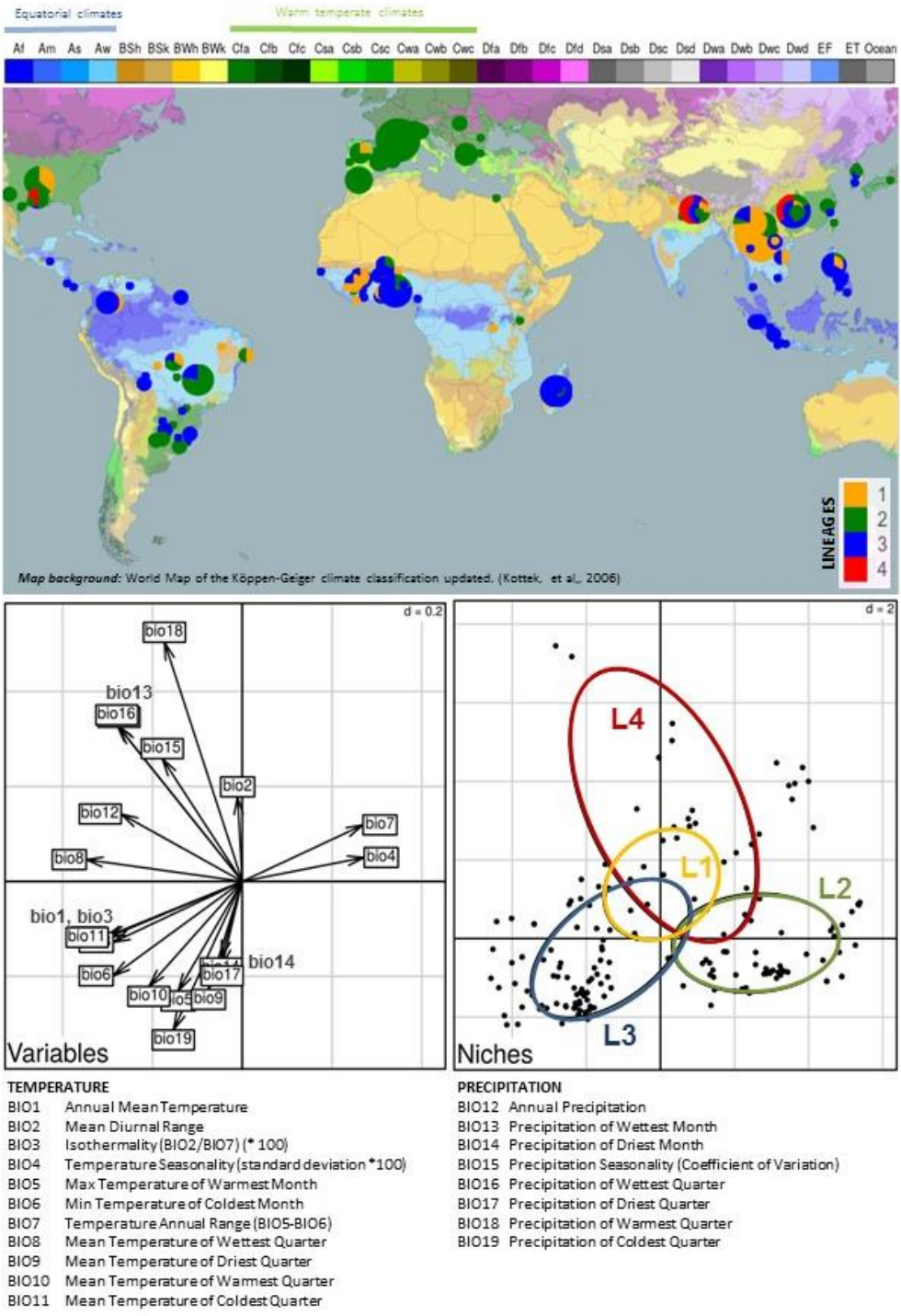


Figure 4: Number of MLG shared between countries (all lineages included).



TEMPERATURE

BIO1	Annual Mean Temperature
BIO2	Mean Diurnal Range
BIO3	Isothermality (BIO2/BIO7) (* 100)
BIO4	Temperature Seasonality (standard deviation *100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO7	Temperature Annual Range (BIO5-BIO6)
BIO8	Mean Temperature of Wettest Quarter
BIO9	Mean Temperature of Driest Quarter
BIO10	Mean Temperature of Warmest Quarter
BIO11	Mean Temperature of Coldest Quarter

PRECIPITATION

BIO12	Annual Precipitation
BIO13	Precipitation of Wettest Month
BIO14	Precipitation of Driest Month
BIO15	Precipitation Seasonality (Coefficient of Variation)
BIO16	Precipitation of Wettest Quarter
BIO17	Precipitation of Driest Quarter
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter

Figure 5: Geographic distribution of four lineages of *P. oryzae* and corresponding climatic data. a: Pie charts representing the distribution of the four lineages, keeping only isolates for which the sampling positions were precisely known (i.e., for which the region, city or GPS position was documented). Background map represents the main climates at world scale as described in Kottek et al. (2006) b and c: OMI method (Outlying Mean Index) to test ecological niches separation of lineages 1, 2, 3 and 4, with b: representing the canonical weights of the environmental variables (bio1 to bio19) and c the site coordinates (dots) and the realized niches of lineages 1, 2 and 3 (circles).

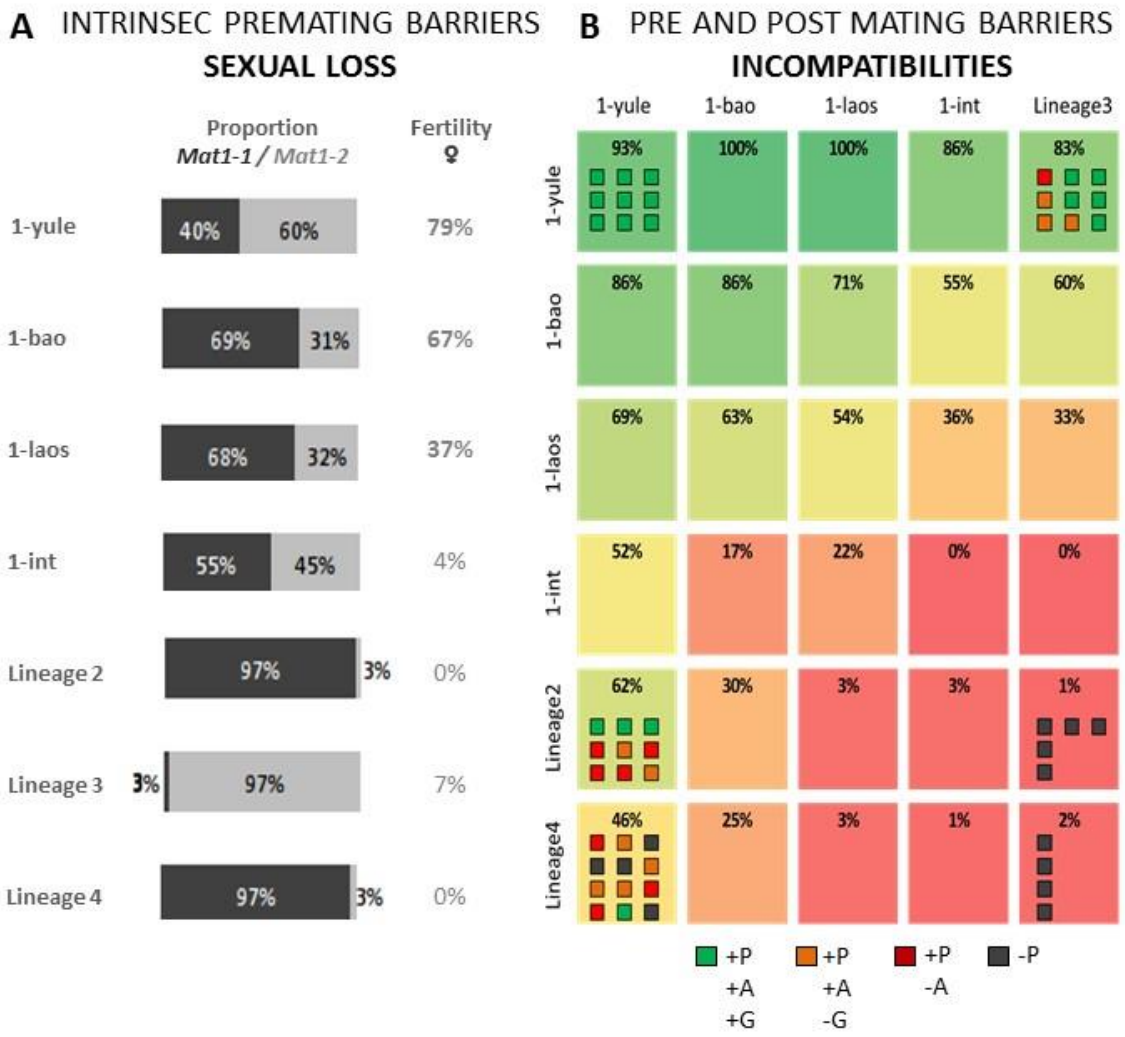


Figure 6: A: Intrinsic ability to sexually reproduce in lineages 2-4 and sub-clusters within lineage 1 of *P. oryzae*: (i) mating-type ratios, (ii) proportion of female fertile isolates; B: Sexual compatibilities between lineages 2-4 and sub-clusters within lineage 1 with (i) proportion of crosses producing at least one perithecium (background color and associated percentage), (ii) scoring asci formation and ascospores germination for a subset of crossings (results are given by the small colored squares, with the following legend: P=perithecia formation, A=asci formation, G=ascospores germination; with + and – indicating cases where P, A and G are positive or negative, respectively).

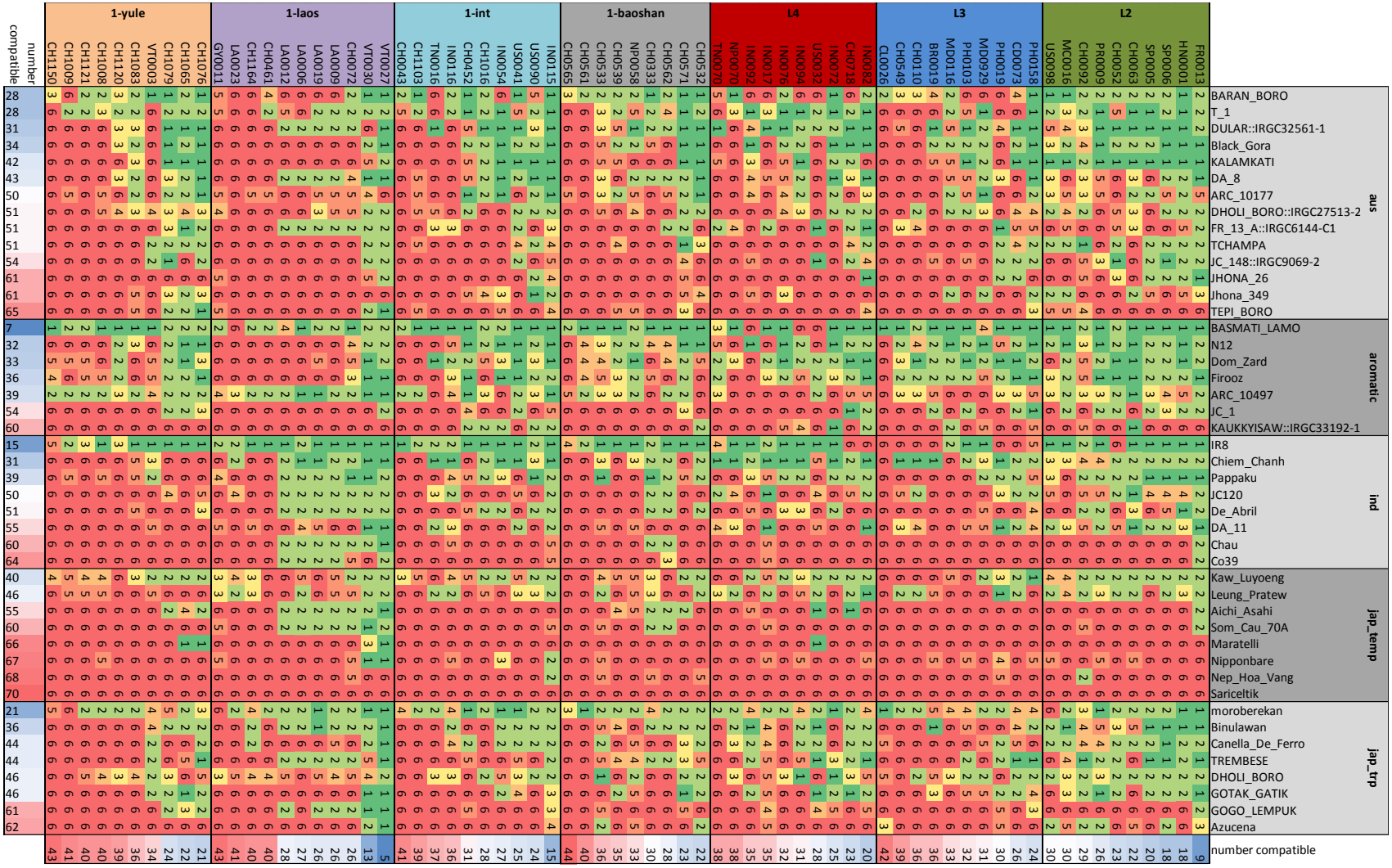


Figure 7: Complete matrix of compatibility between 70 *P. oryzae* isolates and 45 rice plants, as determined in pathogenicity tests in controlled conditions. Numbers in the matrix represent the maximum symptom score across three repetitions. Disease is scored from 1 (no symptoms) to 6 (maximum symptoms) according to Gallet et al., 2016. Numbers at the left and bottom margins represent the number of compatible interactions observed (symptom score >2) for varieties and isolates, respectively.

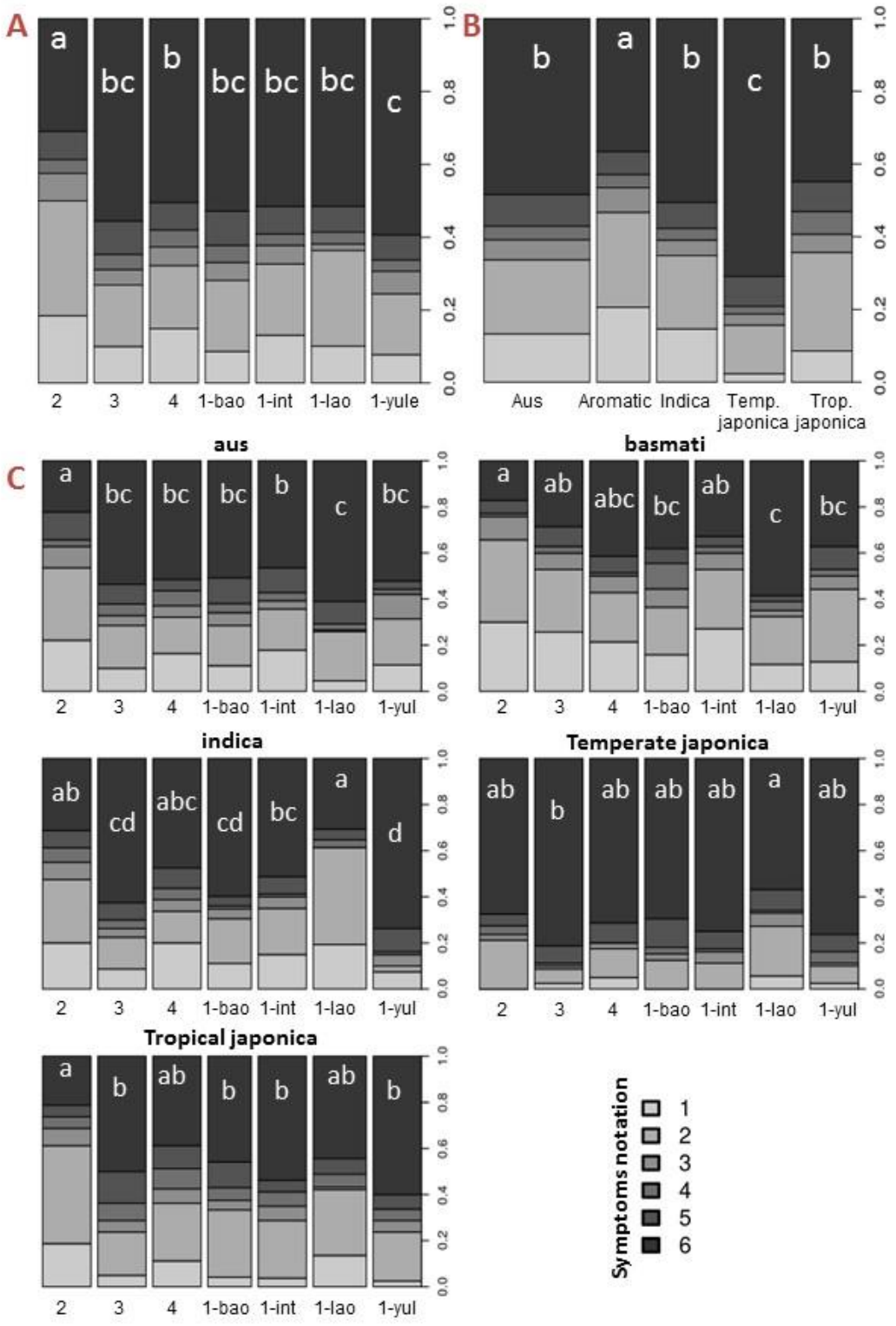


Figure 8: Compatibility between 70 *P. oryzae* isolates and 45 rice varieties, representing five types of rice. A: Proportions of symptom scores as a function of the lineage of origin of isolates; B: Proportions of symptom scores as a function of the type of rice; C: Proportions of symptom as a function of the lineage of origin of isolates, for each type of rice.

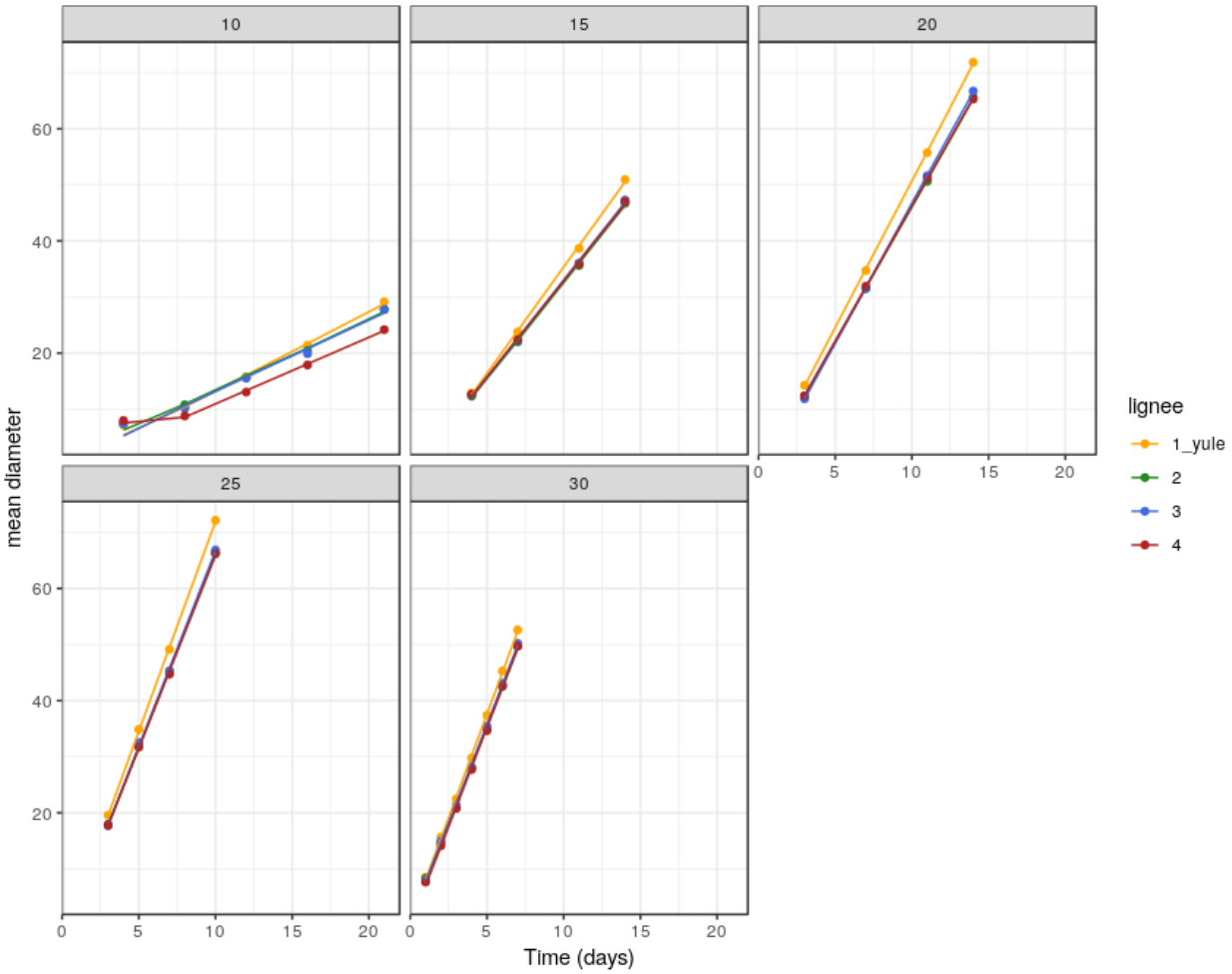


Figure 9: Growth curve for lineages 2-4 and sub-cluster 1-yule within lineage 1 of *P. oryzae* at 5 incubation temperatures.

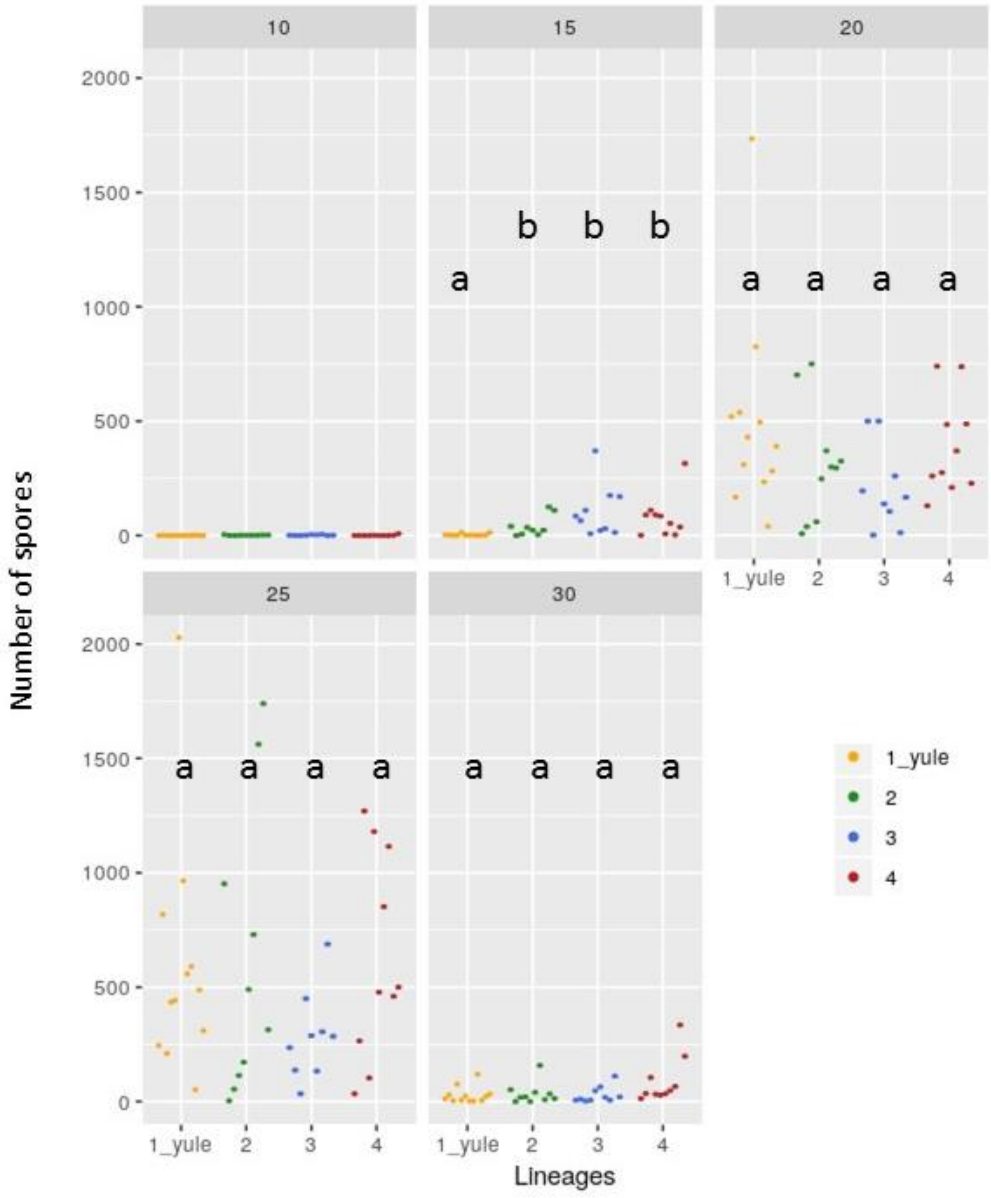
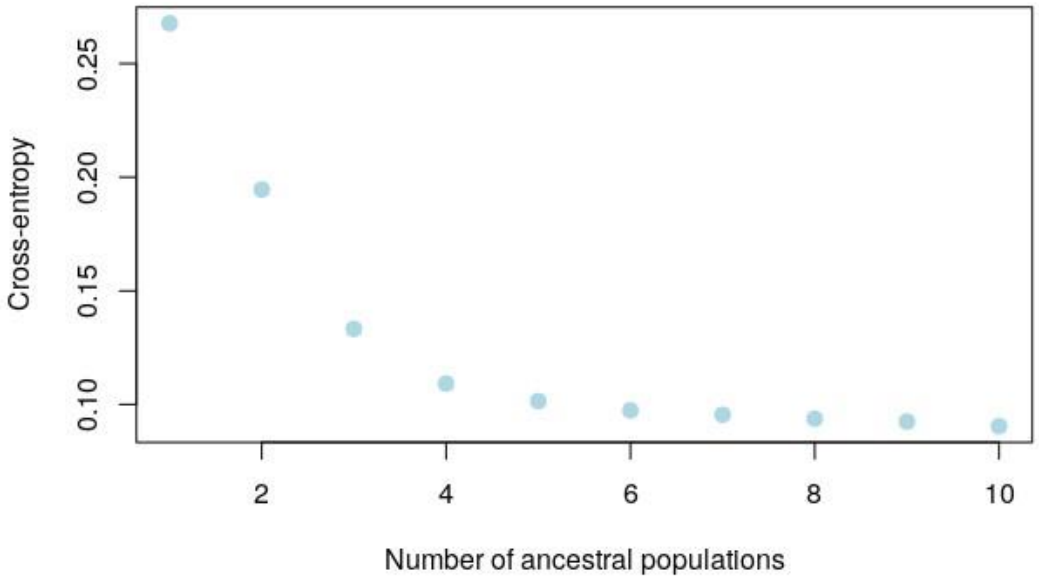


Figure 10: Sporulation rates for lineages 2-4 and sub-cluster 1-yule within lineage 1 of *P. oryzae* at 5 incubation temperatures (average value of three repetitions). Letters indicate significantly differences using a negative binomial generalized linear model.



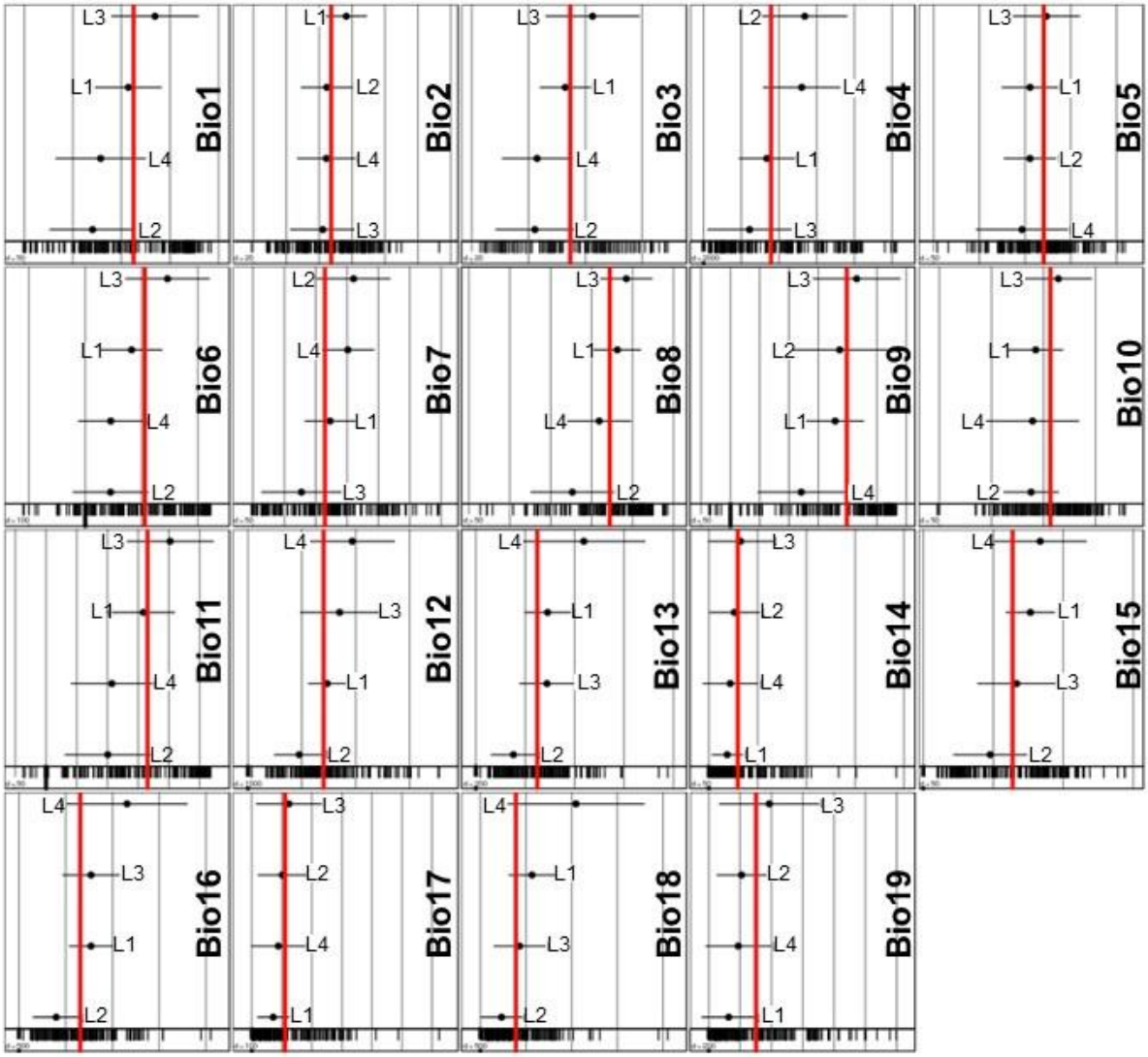
Supplementary figure 1: Cross-entropy (CE) as a function of the number of clusters K modeled in sNMF analyses of population subdivision.

Lineages and L1-subclusters (this study)	Lineages Gladieux et al. 2018						
	1	2	3	4	5	6	Total
1-bao					2		2
1-int						1	1
1-laos	3						3
1-yule	1						1
Lineage2		12					12
Lineage3			14				14
Lineage4				5			5
Total	4	12	14	5	2	1	38

Lineages and L1-subclusters (this study)	World lineages (Saleh et al. 2014)			
	A	B	C	Total
1-bao	7	2		9
1-int	14	2		16
1-laos	17			17
1-yule	55			55
Lineage2	3	69	2	74
Lineage3	3	5	30	38
Lineage4	15	1	2	18
Total	114	79	34	227

Lineages and L1-subclusters (this study)	Asian lineages (Saleh et al. 2014)				
	1	2	3	4	Total
1-bao	5	4			9
1-int	5	1	1	4	11
1-laos	7			10	17
1-yule	5			50	55
Lineage2	1	9	1		11
Lineage3	3	4	10		17
Lineage4	15	1	2		18
Total	41	19	14	64	138

Supplementary figure 2: Assignment of genotypes to clusters identified in this study, in Gladieux et al. 2018b and in Saleh et al. 2014.



Supplementary figure 3: ecological niches of the four major lineages considering each of the 19 biomes individually

Supplementary table 1: FST values between lineage 1 sub-clusters

	1-int	1-laos	1-bao
1-laos	0.1553674		
1-bao	0.2848408	0.4096286	
1-yule	0.2193038	0.180309	0.4867319

Supplementary table 2: List of isolates used in this study

Isolate	Date	Continent	Contry	Province or state	Place (City)	Host genus	Host species	Variety	Mating type	Fertility	Main lineage	L1 subcluster	Clonal groups	Latitude	Longitude	Location accuracy	Phenotyped
AG0003	1997	Americas	Argentina	Corrientes	Mercedes	Oryza	sativa	El Paso 144	2	na	3	-	154	-29.184582	-58.0735719	ville	-
AG0004	1997	Americas	Argentina	Entre Rios	Lucas Norte	Oryza	sativa	Jackson	1	na	2	-	116	-31.4833298	-58.8333282	ville	-
AG0005	1997	Americas	Argentina	Corrientes	Mercedes	Oryza	sativa	Jodon	2	na	3	-	72	-29.184582	-58.0735719	ville	-
AG0007	1996	Americas	Argentina	Entre Rios	Entre Rios	Oryza	sativa	Tacuari	1	na	2	-	116	-31.7746654	-60.4956461	province	-
AG0011	1997	Americas	Argentina	Entre Rios	Lucas Norte	Oryza	sativa	Cnia Mascias	1	na	2	-	116	-31.4833298	-58.8333282	ville	-
AG0012	1997	Americas	Argentina	na	na	Oryza	sativa	Sasaniski	1	na	2	-	116	-38.416097	-63.616672	pays	-
AG0015	1997	Americas	Argentina	Entre Rios	Chajarí	Oryza	sativa	El Paso 144	2	na	3	-	154	-30.7572746	-57.9916191	ville	-
AG0016	1997	Americas	Argentina	Entre Rios	Jubileo	Oryza	sativa	El Paso 144	2	na	3	-	154	-31.7331476	-58.6345258	ville	-
AG0019	na	Americas	Argentina	Corrientes	Rincon de la Luna-Pilagá	Oryza	sativa	El Paso 144	2	na	3	-	154	-28.7	-58.5167	ville	-
AG0020	1998	Americas	Argentina	na	na	Oryza	sativa	Itapé	1	na	2	-	116	-38.416097	-63.616672	pays	-
AG0021	1998	Americas	Argentina	na	na	Oryza	sativa	L425	1	na	2	-	116	-38.416097	-63.616672	pays	-
AG0023	1998	Americas	Argentina	na	na	Oryza	sativa	Colonia Mascias	1	na	2	-	116	-38.416097	-63.616672	pays	-
AG0024	1998	Americas	Argentina	Entre Rios	Entre Rios	Oryza	sativa	Tacuari	1	na	2	-	116	-31.7746654	-60.4956461	province	-
AG0025	1998	Americas	Argentina	na	na	Oryza	sativa	El Paso 144	2	na	3	-	154	-38.416097	-63.616672	pays	-
AG0026	1998	Americas	Argentina	na	na	Oryza	sativa	El Paso 144	2	non-female	2	-	124	-38.416097	-63.616672	pays	-
AG0027	1998	Americas	Argentina	na	na	Oryza	sativa	Chui	2	na	3	-	154	-38.416097	-63.616672	pays	-
AG0036	2000	Americas	Argentina	na	na	Oryza	sativa	Fortuna	1	na	3	-	154	-38.416097	-63.616672	pays	-
AG0037	2000	Americas	Argentina	Chaco	Las Palmas	Oryza	sativa	IR1528-1	na	na	3	-	246	-27.050733	-58.6776801	ville	-
AG0038	2000	Americas	Argentina	na	na	Oryza	sativa	CT 6919	2	na	3	-	154	-38.416097	-63.616672	pays	-
AG0039	2003	Americas	Argentina	na	na	Oryza	sativa	Cambá	2	na	3	-	72	-38.416097	-63.616672	pays	-
AG0042	2004	Americas	Argentina	Corrientes	Virasoro	Oryza	sativa	Fortuna	1	na	2	-	124	-28.0543419	-56.0139884	ville	-
BD0027	1989	Africa	Burundi	na	MURAMBA	Oryza	sativa	Tokambama	1	na	1	Int	254	-2.9351936	30.3594043	ville	-
BF0030	2009	Africa	Burkina Faso	Boulgou	Bagré	Oryza	sativa	FKR19	2	na	3	-	1	11.5156865	-0.542964935	ville	-
BF0031	2009	Africa	Burkina Faso	Boulgou	Bagré	Oryza	sativa	TS2	2	na	3	-	1	11.5156865	-0.542964935	ville	-
BF0034	na	Africa	Burkina Faso	na	Loto	Oryza	sativa	na	na	na	3	-	234	10.9484864	-3.268642784	ville	-
BF0041	na	Africa	Burkina Faso	na	Niéna Dionkéle	Oryza	sativa	v34	na	na	3	-	119	11.7643635	-4.731163198	ville	-
BF0053	na	Africa	Burkina Faso	na	Banfora	Oryza	sativa	B1V1	na	na	3	-	88	10.6312	-4.77876	position gps	-
BF0060	na	Africa	Burkina Faso	na	Banfora	Oryza	sativa	B1V6	na	na	3	-	81	10.6312	-4.77876	position gps	-
BF0061	na	Africa	Burkina Faso	na	Banfora	Oryza	sativa	v140	na	na	3	-	184	10.6312	-4.77876	position gps	-
BF0072	na	Africa	Burkina Faso	na	Pankatioro	Oryza	sativa	na	na	non-female	1	Int	18	10.9115176	-3.902508527	ville	-
BL0003	2009	Americas	Bolivia	Guarayos	San Pablito	Oryza	sativa	Epagri 109	2	na	3	-	85	-15.70028	-63.23859	ville	-
BL0013	2009	Americas	Bolivia	Ichilo	Station Experimental	Oryza	sativa	165	2	na	3	-	72	-17.511027	-63.7695384	province	-
BL0015	2009	Americas	Bolivia	Ichilo	Champ privé	Oryza	sativa	Paititi	2	na	3	-	85	-17.511027	-63.7695384	province	-
BN0013	2011	Africa	Benin	na	Malanville	Oryza	sativa	na	2	non-female	3	-	215	11.8633332	3.3842473	ville	-
BN0019	2011	Africa	Benin	na	Malanville	Oryza	sativa	na	2	non-female	1	Int	18	11.8633332	3.3842473	ville	-
BN0036	2012	Africa	Benin	na	Bétérou	Oryza	sativa	BL19	2	non-female	3	-	77	9.12296	2.16194	position gps	-
BN0047	2012	Africa	Benin	na	Djougou	Oryza	sativa	BL19	1	non-female	2	-	210	9.26512	2.02959	position gps	-
BN0050	2012	Africa	Benin	na	Djougou	Oryza	sativa	BL19	1	non-female	2	-	3	9.26489	2.02973	position gps	-
BN0051	2012	Africa	Benin	na	Kerou	Oryza	sativa	Omihihi	na	na	3	-	81	10.57805	2.12623	position gps	-
BN0070	2012	Africa	Benin	na	Kerou	Oryza	sativa	na	na	na	2	-	239	10.51753	1.06075	position gps	-
BN0083	2012	Africa	Benin	na	Kerou	Oryza	sativa	na	1	non-female	2	-	3	10.51749	1.06078	position gps	-
BN0094	2012	Africa	Benin	na	Kerou	Oryza	sativa	IR841	na	na	3	-	77	10.53916	1.59184	position gps	-
BN0119	2012	Africa	Benin	na	Kokey-Banikoara	Oryza	sativa	Adny11	na	na	4	-	132	11.21857	2.29953	position gps	-
BN0123	2012	Africa	Benin	na	Kokey-Banikoara	Oryza	sativa	Adny11	na	na	4	-	132	11.21851	2.29967	position gps	-
BN0152	2012	Africa	Benin	na	Savalou	Oryza	sp.	wild rice	1	non-female	2	-	210	8.00597	2.04408	position gps	-
BN0155	2012	Africa	Benin	na	Savalou	Oryza	sp.	wild rice	na	na	2	-	210	8.00599	2.04407	position gps	-
BN0157	2012	Africa	Benin	na	Savalou	Oryza	sp.	wild rice	1	non-female	2	-	210	8.00598	2.04409	position gps	-
BN0170	2012	Africa	Benin	na	Kerou	Oryza	sativa	IR841	2	non-female	3	-	23	10.53916	1.59136	position gps	-
BN0188	2012	Africa	Benin	na	Natitingou	Oryza	sativa	Kouantcho	na	na	3	-	23	10.21488	1.2001	position gps	-
BN0192	2010	Africa	Benin	na	Lokossa	Oryza	sativa	DEEJIAOHUALUO	1	non-female	2	-	3	6.6448093	1.7198215	ville	-
BN0202	2010	Africa	Benin	na	Lokossa	Oryza	sativa	CG14	1	non-female	2	-	140	6.6448093	1.7198215	ville	-
BN0256	2010	Africa	Benin	na	Lokossa	Oryza	sativa	DENG PAO ZHAI	na	na	3	-	81	6.6448093	1.7198215	ville	-
BN0307	2013	Africa	Benin	na	Tanguiéta	Oryza	sativa	NERICA	na	na	4	-	132	10.51749	1.1709	position gps	-
BN0342	2013	Africa	Benin	na	Kalalé	Oryza	sativa	local rice	na	na	3	-	81	10.16429	3.22079	position gps	-
BR0003	1985	Americas	Brazil	Goias	Goiana	Oryza	sativa	Moroberekan	1	na	2	-	192	-16.6868912	-49.2647943	ville	-
BR0013	1985	Americas	Brazil	Goias	Goiana	Oryza	sativa	IRAT13	na	na	2	-	150	-16.6868912	-49.2647943	ville	-
BR0019	1986	Americas	Brazil	Goias	Goiana	Oryza	sativa	TETEP	2	non-female	3	-	246	-16.6868912	-49.2647943	ville	yes
BR0049	1990	Americas	Brazil	Mato Grosso	na	Oryza	sativa	IRAT216	1	na	2	-	100	-12.6818712	-56.921099	province	-
BR0054	1990	Americas	Brazil	Mato Grosso	Lucas de rio verde	Oryza	sativa	IRAT216	na	na	2	-	100	-13.064847	-55.9193214	ville	-
BR0059	1990	Americas	Brazil	Alagoas	Coruripe	Oryza	sativa	IAC47	na	na	2	-	117	-10.1250022	-36.1761522	ville	-
BR0060	1990	Americas	Brazil	Alagoas	Coruripe	Oryza	sativa	na	na	non-female	1	Int	2	-10.1250022	-36.1761522	ville	-

BR0061	1990	Americas	Brazil	Piaui	Nozol	Oryza	sativa	METICA1 ET CICA8	na	non-female	1	Int	2	-7.7183401	-42.7289236	province	-
BR0089	1990	Americas	Brazil	Parana	CAMPO MOURAO	Oryza	sativa	na	1	na	2	-	192	-24.0436687	-52.3781045	ville	-
BR0090	1990	Americas	Brazil	Parana	TUPASSI	Oryza	sativa	na	2	na	3	-	246	-24.5883617	-53.5110151	ville	-
BR0099	1989	Americas	Brazil	Goias	Goiania	Oryza	sativa	IAC165	1	na	2	-	237	-16.6868912	-49.2647943	ville	-
BR0100	na	Americas	Brazil	Rondonia	Vilhena	Oryza	sativa	IAC166	na	non-female	1	Int	2	-12.7414031	-60.1304566	ville	-
BR0103	1987	Americas	Brazil	na	Jassiara	Oryza	sativa	Guarani	1	na	2	-	117	-15.9559696	-54.9751847	ville	-
BR0105	1986	Americas	Brazil	Goias	Goiania	Oryza	sativa	Os-6	na	na	2	-	112	-16.6868912	-49.2647943	ville	-
BR0106	1992	Americas	Brazil	Goias	Goiania	Oryza	sativa	CICA 8	na	na	3	-	154	-16.6868912	-49.2647943	ville	-
BR0107	1994	Americas	Brazil	na	na	Oryza	sativa	Aliança	na	na	3	-	154	-14.9872395	-47.30712891	pays	-
BR0108	1994	Americas	Brazil	Goias	Goiania	Oryza	sativa	Metica 1	na	na	2	-	192	-16.6868912	-49.2647943	ville	-
BR0109	1994	Americas	Brazil	Goias	Goiania	Oryza	sativa	Aliança	na	na	3	-	154	-16.6868912	-49.2647943	ville	-
BR0110	1994	Americas	Brazil	Goias	Goiania	Oryza	sativa	CO 39	na	na	2	-	192	-16.6868912	-49.2647943	ville	-
BR0111	1994	Americas	Brazil	Goias	Goiania	Oryza	sativa	C101-A51	na	na	2	-	192	-16.6868912	-49.2647943	ville	-
BR0112	1995	Americas	Brazil	Goias	Goiania	Oryza	sativa	C101LAC	1	na	2	-	51	-16.6868912	-49.2647943	ville	-
BR0114	1995	Americas	Brazil	Goias	Goiania	Oryza	sativa	Mars	na	na	2	-	213	-16.6868912	-49.2647943	ville	-
BR0115	1995	Americas	Brazil	Rio Grande do Sud	Porto Allegre	Oryza	sativa	C105TTP4L23	na	na	3	-	154	-30.0346471	-51.2176584	ville	-
BR0116	1995	Americas	Brazil	Rio Grande do Sud	Porto Allegre	Oryza	sativa	BR IRGA409	na	na	3	-	154	-30.0346471	-51.2176584	ville	-
BR0118	1995	Americas	Brazil	Goias	Goiania	Oryza	sativa	Orizica	na	na	2	-	192	-16.6868912	-49.2647943	ville	-
BR0120	2002	Americas	Brazil	Mato Grosso	Sinop	Oryza	sativa	S2AN488	na	na	2	-	56	-11.8608456	-55.5095451	ville	-
BR0121	2002	Americas	Brazil	Mato Grosso	Sinop	Oryza	sativa	S2AN720	na	non-female	1	Int	2	-11.8608456	-55.5095451	ville	-
BR0122	2002	Americas	Brazil	Mato Grosso	Sinop	Oryza	sativa	S2FD1	na	na	3	-	246	-11.8608456	-55.5095451	ville	-
BR0300	1998	Americas	Brazil	Brasil sud	Don pedrito	Oryza	sativa	El Paso 144	na	na	3	-	246	-14.9872395	-47.30712891	pays	-
BR0304	1999	Americas	Brazil	Rio Grande do Sud	Rio Grande	Oryza	sativa	El Paso 144	na	na	3	-	154	-14.9872395	-47.30712891	pays	-
CD0065	1983	Africa	Ivory coast	na	SAN PEDRO	Oryza	sativa	BKE 189	1	na	1	Int	254	4.7578686	-6.642433	ville	-
CD0073	1981	Africa	ivory coast	na	Bouaké	Oryza	sativa	BP 168	2	non-female	3	-	178	7.690466	-5.0390536	ville	yes
CD0096	1984	Africa	Ivory coast	na	ODIENNE	Oryza	sativa	C 74	2	na	1	Int	120	9.5188754	-7.5572231	ville	-
CD0101	1984	Africa	Ivory coast	na	ODIENNE	Oryza	sativa	BG 90-2	2	na	3	-	203	9.5188754	-7.5572231	ville	-
CD0113	1984	Africa	Ivory coast	na	MAN	Oryza	sativa	IAC 165	na	na	2	-	111	7.4064275	-7.5572231	ville	-
CD0122	1984	Africa	Ivory coast	na	WANINOU	Oryza	sativa	IAC 47	na	non-female	1	Int	254	8.234644	-7.867531	ville	-
CD0142	1989	Africa	Ivory coast	na	ODIENNE	Oryza	sativa	IRAT 216	na	non-female	1	Int	2	9.5188754	-7.5572231	ville	-
CD0203	2003	Africa	Ivory coast	na	Tiassalé (CGRT)	Oryza	sativa	na	2	na	3	-	162	5.9042627	-4.8261424	ville	-
CD0254	2014	Africa	Ivory coast	Tiasssalé	Périmètre de Tiassalé	Oryza	sativa	Wita9	na	na	3	-	113	5.53422	4.49908	position gps	-
CH0043	1991	Asia	China	Yunnan	na	Oryza	sativa	na	2	non-female	1	Int	160	24.4752847	101.3431058	province	yes
CH0051	na	Asia	China	Zhejiang	na	Oryza	sativa	O. Sativa,	1	na	2	-	189	29.1416432	119.7889248	province	-
CH0052	1991	Asia	China	Hunan	na	Oryza	sativa	O. Sativa, B 40	1	non-female	2	-	104	27.6252995	111.8568586	province	yes
CH0063	1985	Asia	China	Hunan	na	Oryza	sativa	O. Sativa	1	non-female	2	-	31	27.6252995	111.8568586	province	yes
CH0072	na	Asia	China	Yunnan	Madjong	Oryza	sativa	na	2	non-female	1	Laos	46	35.86166	104.195397	pays	yes
CH0092	1983	Asia	China	Zhejiang	na	Oryza	sativa	Xiushui 48	1	non-female	2	-	99	29.1416432	119.7889248	province	yes
CH0110	na	Asia	China	Zhejiang	Ruian	Oryza	sativa	Zhufengyuan	2	non-female	3	-	96	27.7778642	120.6560338	ville	yes
CH0328	1998	Asia	China	Hunan	Yanxi IRBN	Oryza	sativa	weiyou 582	na	na	3	-	182	28.127664	110.90203	ville	-
CH0333	1998	Asia	China	Hunan	Yanxi Near-IRBN	Oryza	sativa	Upland rice	1	female	1	Baoshan	201	28.127664	110.90203	ville	yes
CH0337	1998	Asia	China	Hunan	Yanxi IRBN	Oryza	sativa	Guangtaizhan	na	na	3	-	154	28.127664	110.90203	ville	-
CH0338	1998	Asia	China	Hunan	Yanxi IRBN	Oryza	sativa	4620	1	na	2	-	189	28.127664	110.90203	ville	-
CH0452	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	2	non-female	1	Int	89	22.78691	100.977164	ville	yes
CH0454	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	na	1	Laos	76	22.78691	100.977164	ville	-
CH0456	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	2	na	1	Laos	76	22.78691	100.977164	ville	-
CH0460	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	2	na	1	Int	173	22.78691	100.977164	ville	-
CH0461	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	female	1	Laos	76	22.78691	100.977164	ville	yes
CH0462	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	na	1	Laos	76	22.78691	100.977164	ville	-
CH0465	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	na	1	Laos	76	22.78691	100.977164	ville	-
CH0473	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	na	1	Int	255	22.78691	100.977164	ville	-
CH0475	1994	Asia	China	Yunnan	Simao	Oryza	sativa	na	1	na	1	Int	127	22.78691	100.977164	ville	-
CH0531	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	na	1	Baoshan	139	25.112046	99.161761	ville	-
CH0532	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	non-female	1	Baoshan	82	25.112046	99.161761	ville	yes
CH0533	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	non-female	1	Baoshan	228	25.112046	99.161761	ville	yes
CH0539	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	female	1	Baoshan	10	25.112046	99.161761	ville	yes
CH0540	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	na	1	Baoshan	228	25.112046	99.161761	ville	-
CH0549	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	non-female	3	-	233	25.112046	99.161761	ville	yes
CH0550	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	na	3	-	62	25.112046	99.161761	ville	-
CH0551	1997	Asia	China	Yunnan	Yiliang	Oryza	sativa	na	2	na	2	-	7	24.919532	103.141335	ville	-
CH0552	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	na	3	-	233	25.112046	99.161761	ville	-
CH0553	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	na	2	-	8	25.112046	99.161761	ville	-
CH0556	1997	Asia	China	Yunnan	Yiliang	Oryza	sativa	na	1	na	2	-	7	24.919532	103.141335	ville	-

CH0561	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	female	1	Baoshan	61	25.112046	99.161761	ville	yes
CH0562	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	female	1	Baoshan	197	25.112046	99.161761	ville	yes
CH0564	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	na	1	Baoshan	37	25.112046	99.161761	ville	-
CH0565	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	female	1	Baoshan	136	25.112046	99.161761	ville	yes
CH0567	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	1	na	1	Baoshan	228	25.112046	99.161761	ville	-
CH0571	1995	Asia	China	Yunnan	Baoshan	Oryza	sativa	na	2	non-female	1	Baoshan	141	25.112046	99.161761	ville	yes
CH0578	1997	Asia	China	Yunnan	Yiliang	Oryza	sativa	na	1	na	2	-	7	24.919532	103.141335	ville	-
CH0581	1997	Asia	China	Yunnan	Yiliang	Oryza	sativa	na	1	na	2	-	7	24.919532	103.141335	ville	-
CH0595	1997	Asia	China	Yunnan	Yiliang	Oryza	sativa	na	2	na	3	-	96	24.919532	103.141335	ville	-
CH0675	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Zhu 15-Hua 113	na	na	3	-	48	28.1311111	110.8969444	ville	-
CH0676	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Jinyou 974	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0677	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Zhe 733	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0678	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Zhongxian 972	na	na	4	-	202	28.1311111	110.8969444	ville	-
CH0680	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Jinyou 1978	2	na	3	-	81	28.1311111	110.8969444	ville	-
CH0681	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Jinyou 160	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0689	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	T you 291	1	na	2	-	189	28.1311111	110.8969444	ville	-
CH0694	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Xiangzaoxian 31	na	na	4	-	202	28.1311111	110.8969444	ville	-
CH0695	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Lu 18 S/Huachuan	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0696	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	810 S/Huai 96-1	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0701	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	01 Zao-5032	1	na	4	-	202	28.1311111	110.8969444	ville	-
CH0702	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Shanyou 143	1	na	4	-	202	28.1311111	110.8969444	ville	-
CH0704	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Xieyou 80	1	na	2	-	189	28.1311111	110.8969444	ville	-
CH0712	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Yueyou 90	1	na	2	-	189	28.1311111	110.8969444	ville	-
CH0713	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Binyou 0716	1	na	4	-	202	28.1311111	110.8969444	ville	-
CH0714	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Wuyou No.1	na	na	3	-	154	28.1311111	110.8969444	ville	-
CH0718	2001	Asia	China	Hunan	Anhua Yanxi	Oryza	sativa	Jinyou 213	na	non-female	4	-	202	28.1311111	110.8969444	ville	yes
CH0860	2006	Asia	China	Yunnan	Wenshan Dehou	Oryza	sativa	Hexi 35 x rufipogon	1	na	2	-	41	23.3566286	104.2556835	ville	-
CH0866	2006	Asia	China	Yunnan	Wenshan Dehou	Oryza	sativa	Hexi 35 x rufipogon	1	na	2	-	7	23.3566286	104.2556835	ville	-
CH0997	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	female	1	Yule	53	22.0728	100.99431	position gps	yes
CH0998	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	1	na	1	Yule	68	22.0728	100.99431	position gps	-
CH1000	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Int	174	22.0728	100.99431	position gps	-
CH1001	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Yule	212	22.0728	100.99431	position gps	-
CH1002	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Yule	163	22.0728	100.99431	position gps	-
CH1003	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	1	na	1	Yule	207	22.0728	100.99431	position gps	-
CH1004	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Yule	198	22.0728	100.99431	position gps	-
CH1005	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	non-female	1	Yule	253	22.0728	100.99431	position gps	-
CH1007	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	na	non-female	1	Yule	107	22.0728	100.99431	position gps	-
CH1008	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	non-female	1	Yule	217	22.0728	100.99431	position gps	yes
CH1009	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	female	1	Yule	92	22.0728	100.99431	position gps	yes
CH1010	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	na	non-female	1	Yule	107	22.0728	100.99431	position gps	-
CH1011	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Int	110	22.0728	100.99431	position gps	-
CH1012	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	na	non-female	1	Yule	107	22.0728	100.99431	position gps	-
CH1014	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	1	na	1	Yule	208	22.0728	100.99431	position gps	-
CH1016	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	non-female	1	Int	110	22.0728	100.99431	position gps	yes
CH1017	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	1	na	1	Yule	122	22.0728	100.99431	position gps	-
CH1018	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Yule	148	22.0728	100.99431	position gps	-
CH1019	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	2	na	1	Yule	24	22.0728	100.99431	position gps	-
CH1020	2008	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 100	na	non-female	1	Int	193	22.0728	100.99431	position gps	-
CH1064	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	195	22.07277	100.9943	position gps	-
CH1065	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	195	22.07277	100.9943	position gps	yes
CH1066	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	118	22.07277	100.9943	position gps	-
CH1069	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	97	22.07277	100.9943	position gps	-
CH1070	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	86	22.07277	100.9943	position gps	-
CH1071	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	138	22.07277	100.9943	position gps	-
CH1073	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	180	22.07277	100.9943	position gps	-
CH1075	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	175	22.07277	100.9943	position gps	-
CH1076	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	221	22.07277	100.9943	position gps	yes
CH1079	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	40	22.07277	100.9943	position gps	yes
CH1080	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	43	22.07277	100.9943	position gps	-
CH1083	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	38	22.07277	100.9943	position gps	yes
CH1084	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	222	22.07277	100.9943	position gps	-
CH1087	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	29	22.07277	100.9943	position gps	-
CH1088	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	106	22.07277	100.9943	position gps	-

CH1091	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	1	female	1	Yule	118	22.07277	100.9943	position gps	-
CH1092	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	40	22.07277	100.9943	position gps	-
CH1093	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	79	22.07277	100.9943	position gps	-
CH1095	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	71	22.07277	100.9943	position gps	-
CH1096	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	164	22.07277	100.9943	position gps	-
CH1097	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	non-female	1	Yule	168	22.07277	100.9943	position gps	-
CH1098	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 102	2	female	1	Yule	263	22.07277	100.9943	position gps	-
CH1100	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	30	22.07277	100.9943	position gps	-
CH1101	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	249	22.07277	100.9943	position gps	-
CH1103	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	non-female	1	Int	39	22.07277	100.9943	position gps	yes
CH1104	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	non-female	1	Int	128	22.07277	100.9943	position gps	-
CH1105	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	138	22.07277	100.9943	position gps	-
CH1106	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Int	181	22.07277	100.9943	position gps	-
CH1110	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	241	22.07277	100.9943	position gps	-
CH1111	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	75	22.07277	100.9943	position gps	-
CH1112	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	na	non-female	1	Yule	262	22.07277	100.9943	position gps	-
CH1116	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	na	non-female	1	Yule	149	22.07277	100.9943	position gps	-
CH1117	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	22	22.07277	100.9943	position gps	-
CH1118	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Int	245	22.07277	100.9943	position gps	-
CH1119	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	222	22.07277	100.9943	position gps	-
CH1120	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	28	22.07277	100.9943	position gps	yes
CH1121	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	na	non-female	1	Yule	107	22.07277	100.9943	position gps	yes
CH1123	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	219	22.07277	100.9943	position gps	-
CH1124	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	non-female	1	Yule	6	22.07277	100.9943	position gps	-
CH1125	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	2	female	1	Yule	52	22.07277	100.9943	position gps	-
CH1126	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	11	22.07277	100.9943	position gps	-
CH1127	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	145	22.07277	100.9943	position gps	-
CH1130	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 103	1	female	1	Yule	186	22.07277	100.9943	position gps	-
CH1135	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	na	non-female	1	Yule	262	22.07277	100.9943	position gps	-
CH1138	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	na	non-female	1	Yule	107	22.07277	100.9943	position gps	-
CH1140	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	2	female	1	Yule	165	22.07277	100.9943	position gps	-
CH1142	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	1	female	1	Yule	16	22.07277	100.9943	position gps	-
CH1143	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	2	female	1	Yule	108	22.07277	100.9943	position gps	-
CH1144	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	1	female	1	Yule	105	22.07277	100.9943	position gps	-
CH1146	2009	Asia	China	Yunnan	Yule	Oryza	sativa	na	1	female	1	Yule	179	22.07277	100.9943	position gps	-
CH1150	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	2	non-female	1	Yule	27	22.07277	100.9943	position gps	yes
CH1151	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	2	female	1	Yule	147	22.07277	100.9943	position gps	-
CH1152	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	2	female	1	Yule	106	22.07277	100.9943	position gps	-
CH1153	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	1	female	1	Yule	191	22.07277	100.9943	position gps	-
CH1158	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	2	non-female	1	Laos	167	22.07277	100.9943	position gps	-
CH1164	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	1	female	1	Laos	93	22.07277	100.9943	position gps	yes
CH1166	2009	Asia	China	Yunnan	Yule	Oryza	sativa	Yunlu 10	1	female	1	Laos	93	22.07277	100.9943	position gps	-
CL0004	1986	Americas	Colombia	na	Villavicencio	Oryza	sativa	L 201	na	na	3	-	246	4.1373614	-73.62860109	ville	-
CL0006	1986	Americas	Colombia	na	Villavicencio	Oryza	sativa	Metica 1	2	na	3	-	57	4.1373614	-73.62860109	ville	-
CL0008	1986	Americas	Colombia	na	Villavicencio	Oryza	sativa	IR 22	na	na	3	-	246	4.1373614	-73.62860109	ville	-
CL0020	1989	Americas	Colombia	na	META,SANTA ROSA	Oryza	sativa	Metica 1	na	na	2	-	211	4.02477	-73.06698	ville	-
CL0026	1989	Americas	Colombia	na	META,SANTA ROSA	Oryza	sativa	CICA 8	2	non-female	3	-	199	4.02477	-73.06698	ville	yes
CL0029	1989	Americas	Colombia	na	META,SANTA ROSA	Oryza	sativa	CICA 9	na	na	3	-	246	4.02477	-73.06698	ville	-
CL0030	1989	Americas	Colombia	na	META,ALTILLANURA	Oryza	sativa	FANNY	2	na	3	-	72	4.3124959	-72.0789722	ville	-
CL0032	1990	Americas	Colombia	na	META,ALTILLANURA	Oryza	sativa	IAC 165	1	na	2	-	150	4.3124959	-72.0789722	ville	-
CL0033	1990	Americas	Colombia	na	META,ALTILLANURA	Oryza	sativa	CT 8402-6	na	non-female	1	Int	2	4.3124959	-72.0789722	ville	-
CL0035	1990	Americas	Colombia	na	META,ALTILLANURA	Oryza	sativa	CT 6947-7	2	non-female	1	Int	2	4.3124959	-72.0789722	ville	-
CL0067	1995	Americas	Colombia	na	Villavicencio	Oryza	sativa	CICA 9	na	na	3	-	154	4.1373614	-73.62860109	ville	-
CL0079	1996	Americas	Colombia	na	GRANADA	Oryza	sativa	Oryzica llanos 5	na	na	3	-	154	3.544831	-73.705281	ville	-
CL0096	1999	Americas	Colombia	na	na	Oryza	sativa	Ceysvoni	na	na	3	-	78	4.570868	-74.297333	pays	-
CL0099	2000	Americas	Colombia	na	na	Oryza	sativa	Oryzica 1	2	na	3	-	57	4.570868	-74.297333	pays	-
CL0102	2000	Americas	Colombia	na	na	Oryza	sativa	Oryzica 1	2	na	3	-	154	4.570868	-74.297333	pays	-
CL0107	2000	Americas	Colombia	na	na	Oryza	sativa	Oryzica 1	2	na	3	-	57	4.570868	-74.297333	pays	-
CL0115	2000	Americas	Colombia	na	na	Oryza	sativa	Oryzica 1	2	na	3	-	57	4.570868	-74.297333	pays	-
CL0129	2000	Americas	Colombia	na	na	Oryza	sativa	O.caraibe 8	2	na	3	-	78	4.570868	-74.297333	pays	-
CL0133	2000	Americas	Colombia	na	na	Oryza	sativa	O.caraibe 8	2	na	3	-	78	4.570868	-74.297333	pays	-
CL0139	2000	Americas	Colombia	na	na	Oryza	sativa	O.caraibe 8	na	na	3	-	78	4.570868	-74.297333	pays	-
CL0156	2000	Americas	Colombia	na	na	Oryza	sativa	O.caraibe 8	2	na	3	-	78	4.570868	-74.297333	pays	-

CL3.6.7	2000	na	Colombia	Meta	Villavicencio	Oryza	sativa	na	na	na	3	-	125	4.1373614	-73.62860109	ville	-
CM0028	1987	Africa	Cameroon	na	M'BOS	Oryza	sativa	ITA 212	2	na	3	-	203	5.152532	10.0191857	ville	-
CR0002	1986	Asia	Republic of Korea	na	Icheon	Oryza	sativa	Bangambyeo	1	na	2	-	250	37.2081975	127.479663	ville	-
CR0005	1986	Asia	Republic of Korea	na	Suweon	Oryza	sativa	Bonggwangbyeo	1	na	2	-	240	37.2635727	127.0286009	ville	-
CR0008	1990	Asia	Korea	Chungbug	Jaeweon	Oryza	sativa	na	na	na	2	-	250	37.1325821	128.1909481	ville	-
CR0011	1990	Asia	Republic of Korea	Gangweon	Chuncheon	Oryza	sativa	na	na	na	2	-	250	37.8813153	127.7299707	ville	-
CR0033	1985	Asia	Republic of Korea	Jeonnam	Goheung County	Oryza	sativa	na	na	na	3	-	264	34.6112219	127.284978	ville	-
CR0036	1986	Asia	Korea	Chungbug	Ocheon	Oryza	sativa	na	na	na	3	-	157	36.3063646	127.5712809	ville	-
CR0039	na	Asia	Republic of Korea	na	Korea	Oryza	sativa	na	na	na	2	-	250	37.6639976	127.9784585	pays	-
CS0001	2001	Americas	Costa Rica	na	Parrita Mas 1	Oryza	sativa	CR 4338	na	na	3	-	154	9.5201874	-84.3275061	ville	-
EG0015	1997	Africa	Egypt	na	SAKHA	Oryza	sativa	GIZA176	1	na	2	-	134	31.0866656	30.9490733	ville	-
FR0004	1986	Europe	France	Camargue	Arles	Oryza	sativa	RINGO	1	na	2	-	21	43.5939088	4.4689834	ville	-
FR0006	1986	Europe	France	Camargue	Arles	Oryza	sativa	RINGO	1	na	2	-	21	43.5939088	4.4689834	ville	-
FR0013	1988	Europe	France	Camargue	Arles	Oryza	sativa	ROCCA 1	1	non-female	2	-	176	43.5939088	4.4689834	ville	yes
FR0014	1988	Europe	France	Camargue	Arles	Oryza	sativa	ROCCA 2	1	na	2	-	103	43.5939088	4.4689834	ville	-
FR0015	1989	Europe	France	Camargue	Arles	Oryza	sativa	Balila	1	na	2	-	21	43.5939088	4.4689834	ville	-
FR0025	1990	Europe	France	Camargue	Arles	Oryza	sativa	KORAL	1	na	2	-	103	43.5939088	4.4689834	ville	-
FR0038	1993	Europe	France	Camargue	Arles	Oryza	sativa	ARIETE	1	na	2	-	220	43.5939088	4.4689834	ville	-
FR0095	1997	Europe	France	Camargue	Arles	Oryza	sativa	Thaïnato	1	na	2	-	126	43.5939088	4.4689834	ville	-
FR0097	1997	Europe	France	Camargue	Arles	Oryza	sativa	Upla 80	1	na	2	-	32	43.5939088	4.4689834	ville	-
FR0117	2000	Europe	France	Camargue	Arles	Oryza	sativa	Lido	1	na	2	-	134	43.5939088	4.4689834	ville	-
FR0127	2004	Europe	France	Camargue	Arles	Oryza	sativa	Faraman	1	na	2	-	223	43.5939088	4.4689834	ville	-
FR0157	2006	Europe	France	Camargue	Arles	Oryza	sativa	Euro	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0160	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	47	43.5939088	4.4689834	ville	-
FR0167	2006	Europe	France	Camargue	Arles	Oryza	sativa	Euro	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0176	2006	Europe	France	Camargue	Arles	Oryza	sativa	Euro	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0180	2006	Europe	France	Camargue	Arles	Oryza	sativa	Savio	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0181	2006	Europe	France	Camargue	Arles	Oryza	sativa	Savio	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0182	2006	Europe	France	Camargue	Arles	Oryza	sativa	Savio	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0183	2006	Europe	France	Camargue	Arles	Oryza	sativa	Savio	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0185	2006	Europe	France	Camargue	Arles	Oryza	sativa	Savio	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0186	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0193	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	47	43.5939088	4.4689834	ville	-
FR0205	2006	Europe	France	Camargue	Arles	Oryza	sativa	6 SIS 215	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0207	2006	Europe	France	Camargue	Arles	Oryza	sativa	6 SIS 215	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0209	2006	Europe	France	Camargue	Arles	Oryza	sativa	6 SIS 215	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0217	2006	Europe	France	Camargue	Arles	Oryza	sativa	Euro	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0220	2006	Europe	France	Camargue	Arles	Oryza	sativa	TT PR	na	na	2	-	47	43.5939088	4.4689834	ville	-
FR0225	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	80	43.5939088	4.4689834	ville	-
FR0227	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0230	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0233	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0237	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0243	2006	Europe	France	Camargue	Arles	Oryza	sativa	Crodo	na	na	2	-	220	43.5939088	4.4689834	ville	-
FR0255	2007	Europe	France	Camargue	Arles	Oryza	sativa	Lido	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0264	2008	Europe	France	Camargue	Arles	Oryza	sativa	34-88 TM x AAL1	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0314	2009	Europe	France	Camargue	Arles	Oryza	sativa	Ambra	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0321	2009	Europe	France	Camargue	Arles	Oryza	sativa	Sauvage	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0335	2009	Europe	France	Camargue	Arles	Oryza	sativa	Ariete	na	na	2	-	56	43.5939088	4.4689834	ville	-
FR0338	2009	Europe	France	Camargue	Arles	Oryza	sativa	Ariete	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0345	2009	Europe	France	Camargue	Arles	Oryza	sativa	Ariete	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0348	2009	Europe	France	Camargue	Arles	Oryza	sativa	Ariete	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0374	2009	Europe	France	Camargue	Arles	Oryza	sativa	Cultivée	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0678	2012	Europe	France	Camargue	Arles	Oryza	sativa	CRLB1	na	na	2	-	223	43.5939088	4.4689834	ville	-
FR0682	2012	Europe	France	Camargue	Arles	Oryza	sativa	CRLB1	na	na	2	-	126	43.5939088	4.4689834	ville	-
FR0698	2012	Europe	France	Camargue	Arles	Oryza	sativa	Gines 2012	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0699	2012	Europe	France	Camargue	Arles	Oryza	sativa	Gines 2012	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0707	2012	Europe	France	Camargue	Arles	Oryza	sativa	Gines 2012	na	na	2	-	134	43.5939088	4.4689834	ville	-
FR0777	2014	Europe	France	Camargue	Arles	Oryza	sativa	Canarolli	na	na	2	-	56	43.5939088	4.4689834	ville	-
FR0798	2014	Europe	France	Camargue	Arles	Oryza	sativa	Canarolli	na	na	2	-	56	43.5939088	4.4689834	ville	-
FR0924	2014	Europe	France	Camargue	Arles	Oryza	sativa	Caban	na	na	2	-	116	43.5939088	4.4689834	ville	-
GC0001	1990	Africa	Guinea	na	Seredou Macenta	Oryza	sativa	IRAT 216	na	non-female	1	Int	2	8.371919	-9.2868745	ville	-
GH0006	2009	Africa	Ghana	na	Akuse AGRIC farms	Oryza	sativa	Jasmine 85	2	non-female	3	-	81	6.05726	0.08252	position gps	-

GH0007	2012	Africa	Ghana	na	Akuse AGRIC farms	Oryza	sativa	Jasmine 85	2	non-female	3	-	130	6.05726	0.08252	position gps	-
GR0012	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	na	na	2	-	116	40.6310556	22.71283056	position gps	-
GR0013	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	116	40.6310556	22.71283056	position gps	-
GR0014	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	116	40.6310556	22.71283056	position gps	-
GR0016	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	84	40.6310556	22.71283056	position gps	-
GR0017	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	223	40.6310556	22.71283056	position gps	-
GR0027	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Melage	1	na	2	-	126	40.6172389	22.82879722	position gps	-
GR0028	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Melage	1	na	2	-	134	40.6172389	22.82879722	position gps	-
GR0038	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	223	40.6310556	22.70261667	position gps	-
GR0060	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Roma	1	na	2	-	17	41.0512778	23.38379722	position gps	-
GR0062	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Ullisse	1	na	2	-	47	41.0702278	23.37644722	position gps	-
GR0063	2009	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Ullisse	1	na	2	-	223	41.0702278	23.37644722	position gps	-
GR0075	na	Europe	Greece	na	Thessaloniki N.A.G.R.E.F	Oryza	sativa	Gladio	1	na	2	-	259	41.0702278	23.37644722	ville	-
GT0001	1986	Americas	Guatemala	na	IZABAL CRISTINA	Oryza	sativa	DWARF(NAIN)PREC-	2	na	3	-	54	15.3651163	-89.07859667	province	-
GY0011	1978	Americas	French Guyana	na	COMBI	Oryza	sativa	na	2	non-female	1	Laos	95	5.3391018	-52.9419686	ville	yes
GY0028	1999	Americas	French Guyana	na	MANA,COCEROG	Oryza	sativa	BEST 2000 V :	na	na	3	-	72	5.6680095	-53.7789581	ville	-
GY0031	1999	Americas	French Guyana	na	MANA,CROG	Oryza	sativa	CIR	na	na	3	-	72	5.6680095	-53.7789581	ville	-
HN0001	1993	Europe	Hungary	na	Szarvas	Oryza	sativa	OKI-2	1	non-female	2	-	83	46.8635641	20.5526535	ville	yes
HN0006	2008	Europe	Hungary	na	Csarda szallas	Oryza	sativa	SANKA	1	na	2	-	155	46.8919444	20.92527778	position gps	-
HN0008	2008	Europe	Hungary	na	Csarda szallas	Oryza	sativa	SANKA	1	na	2	-	155	46.8919444	20.92527778	position gps	-
HN0018	2008	Europe	Hungary	na	Csarda szallas	Oryza	sativa	M 60	1	na	2	-	155	46.9030556	20.92083333	position gps	-
HN0020	2008	Europe	Hungary	na	Csarda szallas	Oryza	sativa	M 60	1	na	2	-	155	46.9030556	20.92083333	position gps	-
HN0027	2008	Europe	Hungary	na	Kiev jszallas	Oryza	sativa	JANKA	1	na	2	-	87	47.1719444	20.77638889	position gps	-
ID0003	1985	Asia	Indonesia	na	na	Oryza	sativa	na	1	na	2	-	134	-0.789275	113.921327	pays	-
ID0005	1990	Asia	Indonesia	na	Sumatra Sitiung	Oryza	sativa	DINORADO	2	na	3	-	209	-1.0180965	101.6385328	ville	-
ID0013	1991	Asia	Indonesia	na	Sumatra Sitiung	Oryza	sativa	DANAU LAUT	na	na	3	-	226	-1.0180965	101.6385328	ville	-
ID0017	2003	Asia	Indonesia	na	Lampung	Oryza	sativa	Cirata	2	na	3	-	154	-4.5585849	105.4068079	province	-
ID0020	2004	Asia	Indonesia	na	Lampung	Oryza	sativa	Cirata	2	na	3	-	14	-4.5585849	105.4068079	province	-
ID0021	2003	Asia	Indonesia	West Java	Kuningan	Oryza	sativa	Maro	2	na	3	-	154	-7.0138053	108.5700636	ville	-
ID0025	2003	Asia	Indonesia	West Java	Cianjur	Oryza	sativa	IR64	2	na	3	-	225	-6.8168237	107.1425442	ville	-
ID0027	2003	Asia	Indonesia	West Java	Sukabumi	Oryza	sativa	BP206D-KN-51-PN-	2	na	3	-	154	-7.2134052	106.6291304	ville	-
ID0030	2005	Asia	Indonesia	Sud Sumatra	Karang Agung	Oryza	sativa	70215-2-CPA-2-1-B-	2	na	3	-	14	-2.4232741	104.2395942	province	-
ID0190	na	Asia	Indonesia	na	Sumatra Ouest	na	na	CT13432	na	na	3	-	154	-0.9039058	100.652725	province	-
ID0200	na	Asia	Indonesia	na	Sumatra Ouest	na	na	CT13432	na	na	3	-	154	-0.9039058	100.652725	province	-
IN0017	1987	Asia	India	na	na	Oryza	sativa	HR12	1	non-female	4	-	123	20.593685	78.962879	pays	yes
IN0051	1992	Asia	India	na	na	Oryza	sativa	Mt.Vijaya	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0052	1992	Asia	India	na	na	Oryza	sativa	Ponni	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0053	1992	Asia	India	na	na	Oryza	sativa	IR20	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0054	1994	Asia	India	na	na	Oryza	sativa	CO 39	1	non-female	1	Int	58	20.593685	78.962879	pays	yes
IN0055	1994	Asia	India	na	na	Oryza	sativa	Kinadang Patong	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0056	1994	Asia	India	na	na	Oryza	sativa	C101TTP-1	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0057	1992	Asia	India	na	na	Oryza	sativa	Mt.Vijaya	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0058	1992	Asia	India	na	na	Oryza	sativa	Ponni	2	na	3	-	190	20.593685	78.962879	pays	-
IN0059	1992	Asia	India	na	na	Oryza	sativa	Male Sterile Line	1	na	4	-	123	20.593685	78.962879	pays	-
IN0060	1992	Asia	India	na	na	Oryza	sativa	GEB24	1	na	4	-	123	20.593685	78.962879	pays	-
IN0062	1992	Asia	India	na	na	Oryza	sativa	1285	1	na	1	Int	58	20.593685	78.962879	pays	-
IN0066	1992	Asia	India	na	na	Oryza	sativa	Jothi	1	na	1	Int	65	20.593685	78.962879	pays	-
IN0067	1992	Asia	India	na	na	Oryza	sativa	Mukamala	1	na	1	Int	65	20.593685	78.962879	pays	-
IN0069	1993	Asia	India	na	na	Oryza	sativa	IR50	1	na	1	Int	5	20.593685	78.962879	pays	-
IN0071	1992	Asia	India	na	na	Oryza	sativa	IR50	1	na	4	-	205	20.593685	78.962879	pays	-
IN0072	1992	Asia	India	na	na	Oryza	sativa	HR12	1	non-female	4	-	109	20.593685	78.962879	pays	yes
IN0073	1992	Asia	India	na	na	Oryza	sativa	IR50	1	na	4	-	109	20.593685	78.962879	pays	-
IN0074	1992	Asia	India	na	na	Oryza	sativa	Swarna	1	na	4	-	131	20.593685	78.962879	pays	-
IN0075	1993	Asia	India	na	na	Oryza	sativa	IR50	na	na	4	-	205	20.593685	78.962879	pays	-
IN0076	1993	Asia	India	na	na	Oryza	sativa	IR20	1	non-female	4	-	152	20.593685	78.962879	pays	yes
IN0077	1993	Asia	India	na	na	Oryza	sativa	IR20	na	na	4	-	121	20.593685	78.962879	pays	-
IN0078	1993	Asia	India	na	na	Oryza	sativa	ADT36	na	na	4	-	131	20.593685	78.962879	pays	-
IN0079	1994	Asia	India	na	na	Oryza	sativa	IR50	na	na	4	-	131	20.593685	78.962879	pays	-
IN0080	1994	Asia	India	na	na	Oryza	sativa	Fanny	1	na	4	-	121	20.593685	78.962879	pays	-
IN0082	1992	Asia	India	na	na	Oryza	sativa	Jothi	1	non-female	4	-	144	20.593685	78.962879	pays	yes
IN0083	1992	Asia	India	na	na	Oryza	sativa	28.1.1	na	na	4	-	35	20.593685	78.962879	pays	-
IN0084	1992	Asia	India	na	na	Oryza	sativa	M36.1.1	2	na	4	-	115	20.593685	78.962879	pays	-
IN0085	1992	Asia	India	na	na	Oryza	sativa	na	na	na	3	-	48	20.593685	78.962879	pays	-

IN0086	1993	Asia	India	na	na	Oryza	sativa	Ponni	na	na	4	-	35	20.593685	78.962879	pays	-
IN0087	1993	Asia	India	na	na	Oryza	sativa	Bhavani	na	na	4	-	144	20.593685	78.962879	pays	-
IN0089	1994	Asia	India	na	na	Oryza	sativa	Dular	1	na	4	-	132	20.593685	78.962879	pays	-
IN0090	1994	Asia	India	na	na	Oryza	sativa	ASD15	1	na	4	-	35	20.593685	78.962879	pays	-
IN0092	1993	Asia	India	na	na	Oryza	sativa	EK10	1	non-female	4	-	69	20.593685	78.962879	pays	yes
IN0094	1993	Asia	India	na	na	Oryza	sativa	EK10	1	non-female	4	-	242	20.593685	78.962879	pays	yes
IN0096	1994	Asia	India	na	na	Oryza	sativa	OS6	1	na	3	-	154	20.593685	78.962879	pays	-
IN0097	1994	Asia	India	na	na	Oryza	sativa	Yamada Baki	na	na	3	-	154	20.593685	78.962879	pays	-
IN0098	1993	Asia	India	na	na	Oryza	sativa	IR20	2	na	3	-	200	20.593685	78.962879	pays	-
IN0099	1993	Asia	India	na	na	Oryza	sativa	IR20	na	na	3	-	200	20.593685	78.962879	pays	-
IN0102	1993	Asia	India	na	na	Oryza	sativa	CRM41	1	non-female	1	Int	65	20.593685	78.962879	pays	-
IN0104	1993	Asia	India	na	na	Oryza	sativa	Bhavani	2	na	3	-	200	20.593685	78.962879	pays	-
IN0105	1993	Asia	India	na	na	Oryza	sativa	13236	1	na	4	-	202	20.593685	78.962879	pays	-
IN0106	1993	Asia	India	na	na	Oryza	sativa	Lolate	1	na	1	Int	65	20.593685	78.962879	pays	-
IN0107	1993	Asia	India	na	na	Oryza	sativa	CRM41	1	na	1	Int	65	20.593685	78.962879	pays	-
IN0110	1992	Asia	India	Uttar Pradesh	na	Oryza	sativa	na	1	na	2	-	99	26.8467088	80.9461592	province	-
IN0114	1997	Asia	India	Bandrakoti	na	Oryza	sativa	na	2	non-female	1	Int	91	30.5022648	78.3828006	ville	-
IN0115	1997	Asia	India	Matli	na	Oryza	sativa	na	2	non-female	1	Int	12	20.593685	78.962879	pays	yes
IN0116	1997	Asia	India	Matli	na	Oryza	sativa	na	1	non-female	1	Int	90	20.593685	78.962879	pays	yes
IT0002	na	Europe	Italy	Vercelli	na	Oryza	sativa	VERNIA	1	na	2	-	21	45.3202272	8.4185735	ville	-
IT0010	na	Europe	Italy	na	Veneria	Oryza	sativa	IAC 164	1	na	2	-	103	41.87194	12.56738	pays	-
IT0046	na	Europe	Italy	Lombardie	Tavazzano	Oryza	sativa	Albatros	na	na	2	-	220	45.3264129	9.4028281	ville	-
IT0053	na	Europe	Italy	Lombardie	Tavazzano	Oryza	sativa	Aristotele	na	na	2	-	56	45.3264129	9.4028281	ville	-
IT0062	na	Europe	Italy	Lombardie	Tavazzano	Oryza	sativa	Carnise	na	na	2	-	126	45.3264129	9.4028281	ville	-
IT0104	2008	Europe	Italy	Venitie	Verona	Oryza	sativa	Vialone Nana	na	na	2	-	134	45.4383842	10.9916215	ville	-
IT0105	2008	Europe	Italy	Venitie	Verona	Oryza	sativa	Vialone Nana	na	na	2	-	134	45.4383842	10.9916215	ville	-
IT0112	2008	Europe	Italy	Vercelli	Vercelli	Oryza	sativa	Roma	na	na	2	-	134	45.3202272	8.4185735	ville	-
IT0117	2008	Europe	Italy	Vercelli	Vercelli	Oryza	sativa	Volano	na	na	2	-	223	45.3202272	8.4185735	ville	-
IT0135	2009	Europe	Italy	Piémont	Borgolavezzaro	Oryza	sativa	Thaibonnet	na	na	2	-	223	45.3	8.69	position gps	-
IT0138	2009	Europe	Italy	Piémont	Borgolavezzaro	Oryza	sativa	Augusto	na	na	2	-	116	45.3	8.69	position gps	-
IT0140	2009	Europe	Italy	Piémont	Borgolavezzaro	Oryza	sativa	Augusto	na	na	2	-	116	45.3	8.69	position gps	-
IT0143	2009	Europe	Italy	Piémont	Borgolavezzaro	Oryza	sativa	RIS ORT 2	na	na	2	-	220	45.3	8.69	position gps	-
IT0145	2009	Europe	Italy	Piémont	Mortara	Oryza	sativa	Thaibonnet	na	na	2	-	223	45.252204	8.732141	ville	-
JP0003	na	Asia	Japan	na	na	Oryza	sativa	na	2	na	1	Int	98	36.204826	138.252929	pays	-
JP0004	1955	Asia	Japan	na	Aichi Prefecture	Oryza	sativa	na	1	na	2	-	218	35.1801883	136.9065647	province	-
KN0011	2013	Africa	Kenya	na	Mwea	Oryza	sativa	Basmati 370	na	na	2	-	231	-0.6688976	37.3594772	province	-
LA0003	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	2	non-female	1	Laos	64	19.62223	102.055	position gps	-
LA0004	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	2	non-female	1	Laos	177	19.62223	102.055	position gps	-
LA0005	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	66	19.62223	102.055	position gps	-
LA0006	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	114	19.62223	102.055	position gps	yes
LA0007	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	129	19.62223	102.055	position gps	-
LA0009	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	female	1	Laos	216	19.62223	102.055	position gps	yes
LA0010	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	102	19.62223	102.055	position gps	-
LA0012	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	66	19.62223	102.055	position gps	yes
LA0014	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	102	19.62223	102.055	position gps	-
LA0015	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Do Pom	1	non-female	1	Laos	114	19.62223	102.055	position gps	-
LA0018	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	2	female	1	Laos	196	19.62223	102.055	position gps	-
LA0019	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	1	non-female	1	Laos	137	19.62223	102.055	position gps	yes
LA0020	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	2	female	1	Laos	34	19.62223	102.055	position gps	-
LA0021	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	2	female	1	Laos	196	19.62223	102.055	position gps	-
LA0022	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	2	female	1	Laos	44	19.62223	102.055	position gps	-
LA0023	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	1	non-female	1	Laos	251	19.62223	102.055	position gps	yes
LA0024	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	1	female	1	Laos	129	19.62223	102.055	position gps	-
LA0025	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	1	non-female	1	Laos	143	19.62223	102.055	position gps	-
LA0026	2009	Asia	Laos	Luang prabang	Silalek	Oryza	sativa	Khao Mak Khena	1	female	1	Laos	129	19.62223	102.055	position gps	-
MC0004	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	na	na	2	-	116	35.1744271	-6.1473964	ville	-
MC0005	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	na	na	2	-	223	35.1744271	-6.1473964	ville	-
MC0009	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	na	na	2	-	116	35.1744271	-6.1473964	ville	-
MC0013	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	1	na	2	-	223	35.1744271	-6.1473964	ville	-
MC0014	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	1	na	2	-	252	35.1744271	-6.1473964	ville	-
MC0016	1997	Africa	Morocco	na	Tazi	Oryza	sativa	Sariceltik	1	non-female	2	-	135	34.832848	-6.160602	pays	yes
MC0018	1997	Africa	Morocco	na	Larache	Oryza	sativa	V42	na	na	2	-	56	35.1744271	-6.1473964	ville	-
MC0020	1997	Africa	Morocco	na	Larache	Oryza	sativa	Thaibonnet	1	na	2	-	223	35.1744271	-6.1473964	ville	-

MC0021	1997	Africa	Morocco	na	Larache	Oryza	sativa	V53	na	na	2	-	116	35.1744271	-6.1473964	ville	-
MC0041	1999	Africa	Morocco	na	Lahmar Gharb	Oryza	sativa	Lido	1	na	2	-	21	34.729049	-6.2468903	ville	-
MD0116	1992	Africa	Madagascar	na	Sambaina	Oryza	sativa	na	2	female	3	-	185	-18.892453	47.7813995	ville	yes
MD0188	2001	Africa	Madagascar	Lac Alaotra	Ambohitsilaozana	Oryza	sativa	2798	2	na	3	-	170	-17.6993037	48.4674819	ville	-
MD0197	2001	Africa	Madagascar	na	Talata	Oryza	sativa	21	2	na	3	-	185	-18.766949	46.869109	pays	-
MD0202	2001	Africa	Madagascar	na	Talata	Oryza	sativa	39	2	na	3	-	235	-18.766949	46.869109	pays	-
MD0204	2001	Africa	Madagascar	Antsirabe	Tsi-kamo	Oryza	sativa	Manga	na	na	3	-	256	-19.8730077	47.0291162	ville	-
MD0207	2001	Africa	Madagascar	Antsirabe	Antanivao	Oryza	sativa	Manga	na	na	2	-	250	-19.871652	47.035043	ville	-
MD0208	2001	Africa	Madagascar	Antsirabe	Antanivao	Oryza	sativa	na	na	na	2	-	250	-19.871652	47.035043	ville	-
MD0212	2001	Africa	Madagascar	na	Talata	Oryza	sativa	na	na	na	3	-	185	-18.766949	46.869109	pays	-
MD0355	2003	Africa	Madagascar	na	Antananarivo	Oryza	sativa	na	na	na	3	-	170	-18.8791902	47.5079055	ville	-
MD0772	2004	Africa	Madagascar	na	Manandona	Oryza	sativa	Vary Manga	2	na	3	-	170	-20.062611	47.055759	ville	-
MD0790	2004	Africa	Madagascar	na	Moratsiazo	Oryza	sativa	na	2	na	3	-	236	-19.0833333	46.75	ville	-
MD0824	2004	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Exp 208	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD0840	2004	Africa	Madagascar	na	Fifamanor	Oryza	sativa	X265	2	na	3	-	142	-19.8230291	47.1255235	ville	-
MD0866	2004	Africa	Madagascar	na	Fifamanor	Oryza	sativa	Fofifa154	2	na	3	-	36	-19.8230291	47.1255235	ville	-
MD0914	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	36	-19.8041686	47.21424144	ville	-
MD0929	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	female	3	-	232	-19.8041686	47.21424144	ville	yes
MD0931	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD0942	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD0946	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD0948	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD0991	2005	Africa	Madagascar	na	Mangalaza	Oryza	sativa	Fofifa154	2	na	3	-	235	-18.766949	46.869109	pays	-
MD1004	2005	Africa	Madagascar	na	Mangalaza	Oryza	sativa	Fofifa154	2	na	3	-	235	-18.766949	46.869109	pays	-
MD1010	2005	Africa	Madagascar	na	Mangalaza	Oryza	sativa	Fofifa152	2	na	3	-	36	-18.766949	46.869109	pays	-
MD1040	2005	Africa	Madagascar	na	Antananarivo	Oryza	sativa	paysanne	na	na	2	-	250	-18.8791902	47.5079055	ville	-
MD1045	2005	Africa	Madagascar	na	Ambatolampy	Oryza	sativa	Vary Manga	2	na	3	-	185	-19.3836846	47.4392352	province	-
MD1048	2005	Africa	Madagascar	na	Ambatolampy	Oryza	sativa	Vary Manga	2	na	3	-	185	-19.3836846	47.4392352	province	-
MD1054	2005	Africa	Madagascar	na	Manandona	Oryza	sativa	Vary Manga	2	na	3	-	185	-20.062611	47.055759	ville	-
MD1057	2005	Africa	Madagascar	na	Manandona	Oryza	sativa	Vary Manga	2	na	3	-	185	-20.062611	47.055759	ville	-
MD1061	2005	Africa	Madagascar	na	Vinaninkarena	Oryza	sativa	Vary Manga	2	na	3	-	185	4.1373614	-73.62860109	ville	-
MD1080	2005	Africa	Madagascar	na	Manandona	Oryza	sativa	Varimanga	na	na	3	-	185	-20.062611	47.055759	ville	-
MD1096	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD1098	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	36	-19.8041686	47.21424144	ville	-
MD1171	2005	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa161	2	na	3	-	50	-19.8041686	47.21424144	ville	-
MD1231	2006	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa154	2	na	3	-	36	-19.8041686	47.21424144	ville	-
MD1331	2006	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa174	na	na	3	-	50	-19.8041686	47.21424144	ville	-
MD1337	2006	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa180	na	na	3	-	36	-19.8041686	47.21424144	ville	-
MD1341	2006	Africa	Madagascar	na	Andranomanelatra	Oryza	sativa	Fofifa184	na	na	3	-	36	-19.8041686	47.21424144	ville	-
ML0025	1986	Africa	Mali	na	Niema	Oryza	sativa	Sintane Diofor	1	na	3	-	257	12.2166667	-6.616666667	ville	-
ML0060	2009	Africa	Mali	na	Sikasso	Oryza	sativa	IR1529	2	non-female	3	-	203	11.3223834	-5.6983979	ville	-
ML0062	2009	Africa	Mali	na	Sikasso	Oryza	sativa	IRBLKH-K3	1	non-female	1	Int	254	11.3223834	-5.6983979	ville	-
ML0066	2011	Africa	Mali	na	Sikasso	Oryza	sativa	na	1	non-female	1	Int	254	11.3223834	-5.6983979	ville	-
MZ0003	1999	Africa	Mozambique	na	na	Oryza	sativa	na	2	na	3	-	45	-18.665695	35.529562	pays	-
NG0012	2009	Africa	Nigeria	na	Edozhigi	Oryza	sativa	Wita4	2	non-female	3	-	183	9.0763172	5.8775964	ville	-
NG0026	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL12-M	2	non-female	3	-	81	7.3775355	3.9470396	ville	-
NG0041	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL5-M	na	na	3	-	33	7.3775355	3.9470396	ville	-
NG0054	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	CO39	2	non-female	3	-	257	7.3775355	3.9470396	ville	-
NG0065	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBLA-A	na	na	3	-	81	7.3775355	3.9470396	ville	-
NG0095	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBLI-F5	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0102	2010	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBLTA2-PI	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0118	2010	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL9-W	na	na	2	-	7	7.3775355	3.9470396	ville	-
NG0121	2010	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL9-W	1	non-female	2	-	250	7.3775355	3.9470396	ville	-
NG0135	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	Aichi Asahi	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0177	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBLKP-K60	2	non-female	2	-	47	7.3775355	3.9470396	ville	-
NG0190	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL3-CP4	2	non-female	3	-	33	7.3775355	3.9470396	ville	-
NG0191	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL3-CP4	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0192	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL3-CP4	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0197	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL1-CL	2	non-female	3	-	81	7.3775355	3.9470396	ville	-
NG0199	2009	Africa	Nigeria	na	Ibadan	Oryza	sativa	IRBL1-CL	2	non-female	3	-	227	7.3775355	3.9470396	ville	-
NG0201	2010	Africa	Nigeria	na	Ikenne	Oryza	sativa	Aichi Asahi	na	na	3	-	227	6.8717415	3.7105118	ville	-
NG0240	2010	Africa	Nigeria	na	Abéocuta	Oryza	sativa	na	2	non-female	3	-	227	7.147502	3.3619356	ville	-
NG0242	2009	Africa	Nigeria	na	Abéocuta	Oryza	sativa	na	2	non-female	3	-	227	7.147502	3.3619356	ville	-

NG0245	2009	Africa	Nigeria	na	Abéocuta	Oryza	sativa	na	1	na	3	-	33	7.147502	3.3619356	ville	-
NP0004	1989	Asia	Nepal	na	Kathmandu	Oryza	sativa	Pokhrelimasino	2	na	3	-	172	27.7172453	85.3239605	ville	-
NP0012	1992	Asia	Nepal	na	Lumle	Oryza	sativa	Chhoromg	2	na	1	Int	258	28.3762457	83.8313923	ville	-
NP0026	1994	Asia	Nepal	na	Jumla	Oryza	sativa	Sumly Marshi	1	na	1	Int	42	29.2787765	82.12784	province	-
NP0032	2009	Asia	Nepal	Lalitpur	Khumaltar	Oryza	sativa	Khumal 8	na	na	2	-	59	27.6485219	85.3252527	ville	-
NP0036	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0037	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	1	na	4	-	4	28.1739372	84.0973258	province	-
NP0041	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	2	na	3	-	48	28.1739372	84.0973258	province	-
NP0052	2008	Asia	Nepal	Kathmandu	Sangle VDC9	Oryza	sativa	Taichung	1	na	2	-	59	27.800872	85.31943325	ville	-
NP0058	2008	Asia	Nepal	Kathmandu	Sangle VDC9	Oryza	sativa	Taichung 176	1	female	1	Baoshan	101	27.800872	85.31943325	ville	yes
NP0061	2008	Asia	Nepal	Kathmandu	Sangle VDC9	Oryza	sativa	Khumal8	2	na	3	-	172	27.800872	85.31943325	ville	-
NP0068	2009	Asia	Nepal	Kaski	Lumle	Oryza	sativa	local	1	na	4	-	4	28.3762457	83.8313923	ville	-
NP0070	2009	Asia	Nepal	Kaski	Lumle	Oryza	sativa	local	1	non-female	4	-	4	28.3762457	83.8313923	ville	yes
NP0075	2009	Asia	Nepal	Kaski	Lumle	Oryza	sativa	local	1	na	4	-	4	28.3762457	83.8313923	ville	-
NP0077	2009	Asia	Nepal	Kaski	Lumle	Oryza	sativa	local	na	non-female	1	Int	258	28.3762457	83.8313923	ville	-
NP0089	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0091	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0092	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0096	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0103	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	156	28.1739372	84.0973258	province	-
NP0107	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0108	2009	Asia	Nepal	Kaski	Begnas	Oryza	sativa	Masuli	na	na	4	-	4	28.1739372	84.0973258	province	-
NP0129	2008	Asia	Nepal	Kathmandu	Sangle VDC9	Oryza	sativa	Khumal 8	2	na	3	-	172	27.800872	85.31943325	ville	-
NP0138	2008	Asia	Nepal	Kaski	Dhampus-1	Oryza	sativa	Marshi	na	non-female	1	Int	146	28.3068725	83.8419989	ville	-
NP0143	2008	Asia	Nepal	Kaski	Dhampus-1	Oryza	sativa	Marshi	2	na	1	Int	146	28.3068725	83.8419989	ville	-
NP0153	2009	Asia	Nepal	Lalitpur	Godavari	Oryza	sativa	Taichung 176	na	na	2	-	59	27.6086259	85.35933704	ville	-
NP0182	na	Asia	Nepal	na	na	na	na	na	na	na	2	-	59	27.699991	85.333333	pays	-
NP0187	na	Asia	Nepal	na	na	na	na	na	2	na	1	Int	146	27.699991	85.333333	pays	-
NR0001	2013	Africa	Niger	Lossa	Lossa	Oryza	longistaminata	na	na	na	2	-	3	14.22046	1.12394	position gps	-
NR0006	na	Africa	Niger	Bonfeba	Bonfeba	na	na	na	na	na	3	-	184	14.22046	1.12394	position gps	-
NR0009	2013	Africa	Niger	Bonfeba	Bonfeba	Oryza	longistaminata	na	na	na	3	-	184	14.22556	1.11219	position gps	-
NR0024	na	Africa	Niger	Bonfeba	Bonfeba	na	na	na	na	na	3	-	184	14.22046	1.12394	position gps	-
PH0014	1986	Asia	Philippines	na	Los Banos	Oryza	sativa	TETEP	2	na	3	-	190	14.1699121	121.2440631	ville	-
PH0019	1986	Asia	Philippines	na	Los Banos	Oryza	sativa	MILYANG49	2	non-female	3	-	206	14.1699121	121.2440631	ville	yes
PH0091	1983	Asia	Philippines	na	na	Oryza	sativa	DENORADO	2	na	3	-	153	14.599512	120.984219	pays	-
PH0103	na	Asia	Philippines	na	na	Oryza	sativa	na	2	non-female	3	-	214	14.599512	120.984219	pays	yes
PH0118	na	Asia	Philippines	na	na	Oryza	sativa	na	2	na	3	-	260	14.599512	120.984219	pays	-
PH0130	1990	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	IRAT169-F10-6-c Sh	2	na	3	-	154	14.2993252	121.5298637	province	-
PH0143	1992	Asia	Philippines	na	IRRI Blast Nursery	Oryza	sativa	IR36	2	na	3	-	14	14.599512	120.984219	pays	-
PH0151	1990	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	IR55537-13	2	na	3	-	48	14.2993252	121.5298637	province	-
PH0152	1990	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	882110=81-40	na	na	2	-	117	14.2993252	121.5298637	province	-
PH0158	1986	Asia	Philippines	Camarines Sur	IRRI Bo.bula	Oryza	sativa	na	2	non-female	3	-	248	14.599512	120.984219	pays	yes
PH0163	1992	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	IRAT239-1	na	na	3	-	94	14.2993252	121.5298637	province	-
PH0168	1984	Asia	Philippines	na	IRRI	Oryza	sativa	IR29725-3-3-2	2	na	3	-	49	14.599512	120.984219	pays	-
PH0170	1992	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	C101A51	2	na	3	-	133	14.2993252	121.5298637	province	-
PH0173	1984	Asia	Philippines	Batangas	Cuenca	Oryza	sativa	UPLRI-5	2	na	3	-	153	13.9041159	121.0507098	ville	-
PH0178	1992	Asia	Philippines	na	IRRI Caliraya	Oryza	sativa	Fujisaka 5	2	na	3	-	243	14.2993252	121.5298637	province	-
PH0200	2012	Asia	Philippines	Eastern Visayas	Leyte	Oryza	sativa	NSIC 212	2	na	3	-	14	11.3497571	124.4666721	ville	-
PH0211	2012	Asia	Philippines	Luzon	Cavinte, Laguna	Oryza	sativa	Milagrosa	2	non-female	1	Int	2	14.2646732	121.5455292	ville	-
PH0216	2012	Asia	Philippines	Luzon	Cavinte, Laguna	Oryza	sativa	NSIC 222	na	na	3	-	14	14.2646732	121.5455292	ville	-
PH0220	2012	Asia	Philippines	Luzon	Cavinte, Laguna	Oryza	sativa	inconue	2	na	3	-	67	14.2646732	121.5455292	ville	-
PH0242	2012	Asia	Philippines	Negros Island	San Carlos	Oryza	sativa	inconue	2	na	3	-	14	10.476297	123.421711	ville	-
PH0275	2012	Asia	Philippines	Negros Island	San Carlos	Oryza	sativa	152b	2	na	3	-	14	10.476297	123.421711	ville	-
PH0318	2012	Asia	Philippines	Luzon	Nueva Ecija	Oryza	sativa	PYT238	2	na	3	-	14	14.599512	120.984219	pays	-
PH0373	2012	Asia	Philippines	Luzon	Nueva Ecija	Oryza	sativa	PYT 324	2	na	3	-	14	14.599512	120.984219	pays	-
PH0383	2012	Asia	Philippines	Luzon	Nueva Ecija	Oryza	sativa	PR 0? IR 50?	2	na	3	-	14	14.599512	120.984219	pays	-
PH0432	2012	Asia	Philippines	Mindanao	Cotabato	Oryza	sativa	PR37801-15-1-1-3-2-	2	na	3	-	49	7.1083349	125.0388164	province	-
PH0544	2012	Asia	Philippines	Mindanao	Zamboanga del Norte	Oryza	sativa	Bigante	na	na	3	-	94	8.3886282	123.1688883	ville	-
PN0003	1986	Americas	Panama	na	David	Oryza	sativa	Fanny	1	na	3	-	55	8.4007278	-82.4427769	ville	-
PR0009	1991	Europe	Portugal	na	Sado Torre Do Cleri	Oryza	sativa	Ringo	1	non-female	2	-	171	39.72507	-8.999768	pays	yes
PR0025	1991	Europe	Portugal	na	Tejo Coruche Monte V	Oryza	sativa	Thaibonnet	1	na	2	-	223	38.9736645	-8.4403573	ville	-
PR0193	na	Europe	Portugal	na	na	Oryza	sativa	Pegonil 29	na	na	2	-	47	39.72507	-8.999768	pays	-
Pyo_65101022_00012	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-

Pyo_65101022_00016	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65101022_00021	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65101024_00029	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65101104_00048	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	204	13.756331	100.501765	pays	-
Pyo_65101104_00052	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	261	13.756331	100.501765	pays	-
Pyo_65110513_00059	2010	Asia	Thailand	na	na	na	na	RD 23	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65110513_00074	2010	Asia	Thailand	na	na	na	na	RD 15	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65110513_00078	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65110513_00090	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65110513_00094	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65110901_00102	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65111010_00118	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65111010_00130	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65111118_00142	2010	Asia	Thailand	na	na	na	na	NUBN 2	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65111118_00153	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65111118_00158	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65111118_00163	2010	Asia	Thailand	na	na	na	na	RD 15	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65111124_00171	2010	Asia	Thailand	na	na	na	na	LEE NOK	na	na	3	-	154	13.756331	100.501765	pays	-
Pyo_65111124_00176	2010	Asia	Thailand	na	na	na	na	LEE NOK	na	na	3	-	154	13.756331	100.501765	pays	-
Pyo_65111124_00182	2010	Asia	Thailand	na	na	na	na	UBN03-158-KPS-22-	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65111124_00187	2011	Asia	China	na	na	na	na	EK-18 (Inbred)	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111124_00188	2011	Asia	China	na	na	na	na	EK-18 (Inbred)	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111124_00190	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65111124_00192	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111124_00193	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111124_00195	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111124_00200	2011	Asia	China	na	na	na	na	Y liang you 302	na	na	3	-	81	35.86166	104.195397	pays	-
Pyo_65111220_00222	2011	Asia	China	na	na	na	na	na	na	na	3	-	49	35.86166	104.195397	pays	-
Pyo_65111220_00225	2011	Asia	China	na	na	na	na	na	na	na	3	-	49	35.86166	104.195397	pays	-
Pyo_65111220_00227	2011	Asia	China	na	na	na	na	na	na	na	3	-	49	35.86166	104.195397	pays	-
Pyo_65111220_00233	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65111220_00243	2011	Asia	China	na	na	na	na	YiYouHang No.1	na	na	3	-	96	35.86166	104.195397	pays	-
Pyo_65120106_00248	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120106_00272	2011	Asia	China	na	na	na	na	EK-18 (Inbred)	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120106_00277	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120106_00282	2011	Asia	China	na	na	na	na	EK-18 (Inbred)	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120106_00288	2011	Asia	China	na	na	na	na	na	na	na	3	-	49	35.86166	104.195397	pays	-
Pyo_65120106_00292	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120106_00297	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120417_00305	2011	Asia	China	na	na	na	na	na	na	na	3	-	15	35.86166	104.195397	pays	-
Pyo_65120417_00310	2011	Asia	China	na	na	na	na	YiUouHang No.1	na	na	3	-	96	35.86166	104.195397	pays	-
Pyo_65120417_00315	2011	Asia	China	na	na	na	na	Teyou 175	na	na	3	-	81	35.86166	104.195397	pays	-
Pyo_65120417_00320	2011	Asia	China	na	na	na	na	Y liang you 55	na	na	3	-	81	35.86166	104.195397	pays	-
Pyo_65120417_00324	2011	Asia	China	na	na	na	na	Teyou 838	na	na	3	-	81	35.86166	104.195397	pays	-
Pyo_65120424_00356	2011	Asia	China	na	na	na	na	Y liang you 302	na	na	3	-	81	35.86166	104.195397	pays	-
Pyo_65120424_00365	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00370	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00375	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00380	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00385	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120424_00386	2011	Asia	China	na	na	na	na	na	na	na	3	-	182	35.86166	104.195397	pays	-
Pyo_65120424_00390	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00394	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120424_00397	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120502_00399	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120502_00401	2011	Asia	China	na	na	na	na	na	na	na	3	-	154	35.86166	104.195397	pays	-
Pyo_65120716_00405	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	14	11.6522364	108.5229492	pays	-
Pyo_65120716_00409	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	14	11.6522364	108.5229492	pays	-
Pyo_65120716_00417	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	14	11.6522364	108.5229492	pays	-
Pyo_65120716_00423	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	14	11.6522364	108.5229492	pays	-
Pyo_65120719_00429	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	14	11.6522364	108.5229492	pays	-
Pyo_65120719_00431	2012	Asia	Indonesia	na	na	na	na	Ciherang (local	na	na	3	-	154	-0.789275	113.921327	pays	-
Pyo_65120719_00434	2012	Asia	Indonesia	na	na	na	na	na	na	na	3	-	154	-0.789275	113.921327	pays	-

Pyo_65120719_00439	2012	Asia	Indonesia	na	na	na	na	Inpari 10 (local	na	na	3	-	225	-0.789275	113.921327	pays	-
Pyo_65120719_00445	2012	Asia	Indonesia	na	na	na	na	Inpari 10 (local	na	na	3	-	154	-0.789275	113.921327	pays	-
Pyo_65120723_00453	2010	Asia	Thailand	na	na	na	na	UBN 04128-116	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120723_00458	2010	Asia	Thailand	na	na	na	na	UBN 04128-119	na	na	3	-	261	13.756331	100.501765	pays	-
Pyo_65120723_00459	2010	Asia	Thailand	na	na	na	na	RD 15	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120723_00464	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120723_00470	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120806_00475	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120808_00480	2010	Asia	Thailand	na	na	na	na	UBN 04128-119	na	na	3	-	261	13.756331	100.501765	pays	-
Pyo_65120808_00484	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120914_00488	2010	Asia	Thailand	na	na	na	na	CHAINAT 1	na	na	3	-	94	13.756331	100.501765	pays	-
Pyo_65120914_00495	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120914_00496	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65120914_00500	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00502	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00503	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00511	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00517	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00518	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120921_00521	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120926_00527	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120928_00528	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120928_00532	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120928_00533	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120928_00536	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65120928_00541	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65120928_00545	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121001_00548	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00549	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00555	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00557	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00562	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121003_00568	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00569	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00573	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00574	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121003_00578	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121003_00583	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121126_00590	2010	Asia	Thailand	na	na	na	na	UBN03-164-KPS-29-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121126_00598	2010	Asia	Thailand	na	na	na	na	UBN0429-2-1-R-2	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121126_00603	2010	Asia	Thailand	na	na	na	na	UBN04029-2-1-R-1	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121126_00608	2010	Asia	Thailand	na	na	na	na	UBN02-248-9-3-	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121126_00612	2010	Asia	Thailand	na	na	na	na	HANG-NAK	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121126_00615	2010	Asia	Thailand	na	na	na	na	UBN02-248-9-3-	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121126_00619	2010	Asia	Thailand	na	na	na	na	UBN04050-22-2-R-1	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121126_00625	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121126_00626	2010	Asia	Thailand	na	na	na	na	KKN-012021-43-2-1-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121126_00630	2010	Asia	Thailand	na	na	na	na	NANG-LUEN	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121205_00631	2010	Asia	Thailand	na	na	na	na	KDML 105 (LOCAL)	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121205_00633	2010	Asia	Thailand	na	na	na	na	KKN97062-B-6-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121205_00638	2010	Asia	Thailand	na	na	na	na	UBN02-248-9-3-	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121205_00641	2010	Asia	Thailand	na	na	na	na	UBN04050-22-2-R-1	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00642	2010	Asia	Thailand	na	na	na	na	KKN97062-B-6-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00643	2010	Asia	Thailand	na	na	na	na	UBN04016-16-1-R-1	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00644	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00645	2010	Asia	Thailand	na	na	na	na	KKN97062-B-6-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00646	2010	Asia	Thailand	na	na	na	na	KKN97038-39-1-2-1-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00647	2010	Asia	Thailand	na	na	na	na	KEE-TOM-YAI	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121210_00652	2010	Asia	Thailand	na	na	na	na	YUAN-KLANG	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00657	2010	Asia	Thailand	na	na	na	na	KKN04023-NK1-8-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00662	2010	Asia	Thailand	na	na	na	na	YUAN-KLANG	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121210_00664	2010	Asia	Thailand	na	na	na	na	UBN02-248-44-3-	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121210_00669	2010	Asia	Thailand	na	na	na	na	UBN021277-3-NRM-	na	na	3	-	60	13.756331	100.501765	pays	-

Pyo_65121210_00673	2010	Asia	Thailand	na	na	na	na	UBN04050-22-3-R-3	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121210_00678	2010	Asia	Thailand	na	na	na	na	Niew-Dam	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121218_00684	2010	Asia	Thailand	na	na	na	na	UBN02-248-9-3-	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65121218_00685	2010	Asia	Thailand	na	na	na	na	UBN02-248-9-3-	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121218_00686	2010	Asia	Thailand	na	na	na	na	UBN04050-22-2-R-1	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121218_00687	2010	Asia	Thailand	na	na	na	na	HANG-NAK	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121218_00689	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65121218_00690	2010	Asia	Thailand	na	na	na	na	KKN97062-B-6-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121218_00691	2010	Asia	Thailand	na	na	na	na	NANG-LUEN	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65121221_00692	2010	Asia	Thailand	na	na	na	na	YUAN-KLANG	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65130311_00693	2010	Asia	Thailand	na	na	na	na	UBN060061-4-15-3-	na	na	3	-	151	13.756331	100.501765	pays	-
Pyo_65130311_00699	2010	Asia	Thailand	na	na	na	na	UBN04050-22-3-R-5	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130311_00705	2010	Asia	Thailand	na	na	na	na	UBN04029-2-3-R-4	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130311_00711	2010	Asia	Thailand	na	na	na	na	KKN98018-43-1-2-3	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65130311_00715	2010	Asia	Thailand	na	na	na	na	KKN04023-NK1-8-2-	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65130311_00721	2010	Asia	Thailand	na	na	na	na	RD 10	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65130312_00726	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00727	2010	Asia	Thailand	na	na	na	na	RA CHI NI 2	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00732	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00736	2010	Asia	Thailand	na	na	na	na	RD 15	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00741	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	70	13.756331	100.501765	pays	-
Pyo_65130312_00746	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00751	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130312_00757	2010	Asia	Thailand	na	na	na	na	UBN03-82-KPS-9-9	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130325_00759	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130424_00761	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130503_00762	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130503_00767	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130503_00772	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130503_00774	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130503_00777	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130503_00782	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130503_00787	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130503_00792	2010	Asia	Thailand	na	na	na	na	RD 6	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130503_00797	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	60	13.756331	100.501765	pays	-
Pyo_65130503_00802	2010	Asia	Thailand	na	na	na	na	CHAINAT 1	na	na	3	-	94	13.756331	100.501765	pays	-
Pyo_65130515_00804	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00807	2010	Asia	Thailand	na	na	na	na	na	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00812	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00817	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00820	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00825	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00830	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130515_00835	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00836	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00841	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00846	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00851	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00856	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00861	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130702_00866	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130704_00871	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130704_00876	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130704_00886	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130704_00890	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130704_00895	2010	Asia	Thailand	na	na	na	na	Neiw Dum	na	na	3	-	25	13.756331	100.501765	pays	-
Pyo_65130704_00900	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130711_00904	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130711_00906	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130719_00907	2013	Americas	Brazil	na	na	na	na	Inov (hibrid)	na	na	3	-	96	-14.9872395	-47.30712891	pays	-
Pyo_65130719_00912	2013	Americas	Brazil	na	na	na	na	Puitá Inta CL	na	na	3	-	96	-14.9872395	-47.30712891	pays	-
Pyo_65130719_00917	2013	Americas	Brazil	na	na	na	na	Puitá Inta CL	na	na	3	-	96	-14.9872395	-47.30712891	pays	-
Pyo_65130719_00921	2013	Americas	Brazil	na	na	na	na	Puitá Inta CL	na	na	3	-	96	-14.9872395	-47.30712891	pays	-

Pyo_65130719_00926	2013	Americas	Brazil	na	na	na	na	Puitá Inta CL	na	na	3	-	246	-14.9872395	-47.30712891	pays	-
Pyo_65130719_00927	2013	Americas	Brazil	na	na	na	na	Delotas	na	na	3	-	96	-14.9872395	-47.30712891	pays	-
Pyo_65130719_00937	2010	Asia	Thailand	na	na	na	na	KDML 105	na	na	3	-	200	13.756331	100.501765	pays	-
Pyo_65130723_00938	2013	Americas	Brazil	na	na	na	na	Puitá Inta CL	na	na	3	-	96	-14.9872395	-47.30712891	pays	-
Pyo_65131219_00939	2013	Asia	Philippines	na	na	na	na	RC220	na	na	3	-	14	14.599512	120.984219	pays	-
Pyo_65131219_00944	2012	Asia	Vietnam	na	na	na	na	Khan dan 18	na	na	3	-	154	11.6522364	108.5229492	pays	-
Pyo_65131219_00949	2012	Asia	Vietnam	na	na	na	na	OM 10383	na	na	3	-	230	11.6522364	108.5229492	pays	-
Pyo_65131219_00953	2012	Asia	Vietnam	na	na	na	na	OM 6976	na	na	3	-	49	11.6522364	108.5229492	pays	-
Pyo_65150527_00963	2014	Asia	Vietnam	na	na	na	na	B-TE1	na	na	3	-	182	11.6522364	108.5229492	pays	-
Pyo_65150527_00965	2014	Asia	Vietnam	na	na	na	na	Thao Duoc	na	na	3	-	63	11.6522364	108.5229492	pays	-
Pyo_65150527_00966	2014	Asia	Vietnam	na	na	na	na	LT2	na	na	3	-	154	11.6522364	108.5229492	pays	-
Pyo_65150527_00967	2014	Asia	Vietnam	na	na	na	na	BC-15	na	na	3	-	182	11.6522364	108.5229492	pays	-
Pyo_65150527_00969	2014	Asia	Vietnam	na	na	na	na	BC-15 (No 3)	na	na	3	-	154	11.6522364	108.5229492	pays	-
Pyo_65150527_00970	2014	Asia	Vietnam	na	na	na	na	BC-15 (No 3)	na	na	3	-	154	11.6522364	108.5229492	pays	-
Pyo_65150527_00972	2015	Asia	Vietnam	na	na	na	na	AGPPS135	na	na	3	-	63	11.6522364	108.5229492	pays	-
Pyo_65150527_00973	2015	Asia	Vietnam	na	na	na	na	OM10447	na	na	3	-	63	11.6522364	108.5229492	pays	-
Pyo_65150715_00975	2015	Asia	Vietnam	na	na	na	na	OM6976	na	na	3	-	49	11.6522364	108.5229492	pays	-
Pyo_65150715_00976	2013	Asia	Bangladesh	na	na	na	na	BRR1 Dhan-34	na	na	4	-	238	23.684994	90.356331	pays	-
Pyo_65150715_00977	2013	Asia	Bangladesh	na	na	na	na	BRR1 Dhan-34	na	na	4	-	238	23.684994	90.356331	pays	-
Pyo_65150715_00978	2013	Asia	Bangladesh	na	na	na	na	BRR1 Dhan-34	na	na	4	-	238	23.684994	90.356331	pays	-
Pyo_65150715_00979	2013	Asia	Bangladesh	na	na	na	na	BRR1 Dhan-34	na	na	4	-	238	23.684994	90.356331	pays	-
RM0001	1997	Europe	Roumania	na	Oltenita	Oryza	sativa	Polizesti	1	na	2	-	188	44.0854664	26.637483	ville	-
RW0010	1990	Africa	Rwanda	na	Buzeni	Oryza	sativa	ROJOFOTSY (1285)	1	na	1	Laos	166	-1.940278	29.873888	pays	-
SN0009	1986	Africa	Senegal	na	Djibelor	Oryza	sativa	BARAFITA	2	na	3	-	81	12.552757	-16.322291	ville	-
SP0005	1994	Europe	Spain	Tarragone	AMPOSTA,SALATS,DELTA	Oryza	sativa	LIDO	1	non-female	2	-	159	40.6837192	0.6298922	ville	yes
SP0006	1994	Europe	Spain	Aragon	na	Oryza	sativa	HUESCA	1	non-female	2	-	126	41.5976275	-0.9056623	province	yes
SP0022	1999	Europe	Spain	SEVILLA	Villafranco del Guadalquivir	Oryza	sativa	HISPAGRAN	na	na	2	-	126	37.132122	-6.165179	ville	-
SP0026	2000	Europe	Spain	Andalusia	CIFA-Las Torres	Oryza	sativa	Puntal A5	na	na	2	-	116	40.9080445	-5.6530252	ville	-
SP0038	2000	Europe	Spain	Andalusia	CIFA-Las Torres	Oryza	sativa	Puntal A12	na	non-female	1	Int	98	40.9080445	-5.6530252	ville	-
SP0043	2000	Europe	Spain	Andalusia	CIFA-Las Torres	Oryza	sativa	Puntal A18	na	na	2	-	223	40.9080445	-5.6530252	ville	-
SP0044	2000	Europe	Spain	Andalusia	CIFA-Las Torres	Oryza	sativa	Puntal A18	na	na	2	-	223	40.9080445	-5.6530252	ville	-
SP0125	2006	Europe	Spain	na	Silla	Oryza	sativa	Montsianell	na	na	2	-	56	39.3626846	-0.4120058	ville	-
SP0141	2006	Europe	Spain	Tarragone	na	Oryza	sativa	Baixet	na	na	2	-	126	41.1188827	1.2444909	province	-
SP0143	2006	Europe	Spain	Tarragone	na	Oryza	sativa	Baixet	na	na	2	-	126	41.1188827	1.2444909	province	-
SP0149	2006	Europe	Spain	Tarragone	na	Oryza	sativa	Sarlet	1	na	2	-	126	41.1188827	1.2444909	province	-
SP0162	2006	Europe	Spain	Tarragone	na	Oryza	sativa	Palacios	na	na	2	-	126	41.1188827	1.2444909	province	-
SP0167	2006	Europe	Spain	na	Pego (junta arroz ecologico)	Oryza	sativa	Bomba	na	na	2	-	56	38.8429246	-0.1156943	ville	-
SP0168	2006	Europe	Spain	na	Casa catalans (sollana)	Oryza	sativa	Bomba	na	na	2	-	56	39.2779662	-0.3830231	ville	-
SP0172	2006	Europe	Spain	na	Cavallo de la llonga (llonga)	Oryza	sativa	Bomba	na	na	2	-	56	39.368948	3.2213897	ville	-
SP0173	2006	Europe	Spain	na	Campillo-casa-cubella	Oryza	sativa	Senia	na	na	2	-	220	39.2779662	-0.3830231	ville	-
SP0179	2006	Europe	Spain	na	Compra (sueca)	Oryza	sativa	Gleva	na	na	2	-	220	39.2033565	-0.3113278	ville	-
SP0184	2006	Europe	Spain	na	Cami sequita nova (llonga)	Oryza	sativa	Bomba	na	na	2	-	56	39.368948	3.2213897	ville	-
SP0216	2007	Europe	Spain	Valance	Sollana	Oryza	sativa	Bomba	na	na	2	-	56	39.2779662	-0.3830231	ville	-
SP0227	2007	Europe	Spain	Valance	Sollana	Oryza	sativa	Bomba	na	na	2	-	56	39.2779662	-0.3830231	ville	-
TG0004	2012	Africa	Togo	na	Elavayo	Oryza	sativa	Nérica L19	1	non-female	1	Int	58	7.15421	0.4249	position gps	-
TG0005	2012	Africa	Togo	na	Assomè	Oryza	sativa	Nérica L19	2	non-female	3	-	130	6.21265	1.17937	position gps	-
TG0008	2012	Africa	Togo	na	Elavayo	Oryza	sativa	Nérica L19	1	non-female	1	Int	58	7.15405	0.42466	position gps	-
TG0011	2012	Africa	Togo	na	Elavayo	Oryza	sativa	Nérica L19	1	non-female	1	Int	58	7.15422	0.42489	position gps	-
TG0032	2012	Africa	Togo	na	Elavayo	Oryza	sativa	Nérica L19	1	non-female	1	Int	58	7.15422	0.42489	position gps	-
TG0043	2012	Africa	Togo	na	Elavayo	Oryza	sativa	Nérica L19	1	non-female	1	Int	58	7.15405	0.42466	position gps	-
TH0001	1980	Asia	Thailand	na	Swanfarm centre	Oryza	sativa	na	2	na	3	-	261	13.756331	100.501765	pays	-
TH0004	1987	Asia	Thailand	na	Chengmai	Oryza	sativa	MAECHUN	na	na	3	-	9	18.787747	98.9931284	ville	-
TH0007	1980	Asia	Thailand	na	Thammrat sud	Oryza	sativa	HY 71	na	na	3	-	261	13.756331	100.501765	pays	-
TH0052	na	Asia	Thailand	na	Nakhonsi	Oryza	sativa	na	2	na	3	-	244	8.4303975	99.9631219	province	-
TN0001	2013	Africa	Tanzania	na	na	Oryza	sativa	na	1	non-female	1	Int	187	-6.369028	34.888822	pays	-
TN0002	2013	Africa	Tanzania	na	na	Oryza	sativa	na	na	non-female	1	Int	187	-6.369028	34.888822	pays	-
TN0014	2013	Africa	Tanzania	na	na	Oryza	sativa	Zambia	1	non-female	1	Int	187	-6.369028	34.888822	pays	-
TN0015	2013	Africa	Tanzania	na	na	Oryza	sativa	Zambia	1	non-female	1	Int	194	-6.369028	34.888822	pays	-
TN0016	2013	Africa	Tanzania	na	na	Oryza	sativa	Zambia	1	non-female	1	Int	187	-6.369028	34.888822	pays	yes
TN0038	2013	Africa	Tanzania	na	na	Oryza	sativa	na	1	non-female	4	-	20	-6.369028	34.888822	pays	-
TN0042	2013	Africa	Tanzania	na	na	Oryza	sativa	na	na	na	4	-	20	-6.369028	34.888822	pays	-
TN0045	2013	Africa	Tanzania	na	na	Oryza	sativa	na	1	non-female	4	-	20	-6.369028	34.888822	pays	-
TN0057	2013	Africa	Tanzania	na	na	Oryza	sativa	Supa	1	non-female	4	-	20	-6.369028	34.888822	pays	-

TN0068	2013	Africa	Tanzania	na	na	Oryza	sativa	Supa	1	non-female	4	-	20	-6.369028	34.888822	pays	-
TN0070	2013	Africa	Tanzania	na	na	Oryza	sativa	Supa	1	non-female	4	-	20	-6.369028	34.888822	pays	yes
TN0078	2013	Africa	Tanzania	na	na	Oryza	sativa	Supa	1	non-female	4	-	20	-6.369028	34.888822	pays	-
TN0090	2013	Africa	Tanzania	na	na	Oryza	sativa	na	1	non-female	1	Int	187	-6.369028	34.888822	pays	-
TR0006	1995	Europe	Turkey	na	Balabanköy Uzunköprü	Oryza	sativa	Serhat-92	1	na	2	-	21	41.090075	26.549133	ville	-
TR0025	na	Europe	Turkey	na	na	Oryza	sativa	Baldo	na	na	2	-	21	38.963745	35.243322	pays	-
TR0032	na	Europe	Turkey	na	na	Oryza	sativa	Baldo	na	na	2	-	21	38.963745	35.243322	pays	-
TR0033	na	Europe	Turkey	na	na	Oryza	sativa	Baldo	na	na	2	-	21	38.963745	35.243322	pays	-
TR0036	na	Europe	Turkey	na	na	Oryza	sativa	Baldo	na	na	2	-	21	38.963745	35.243322	pays	-
TW0001	1980	Asia	Taiwan	na	Taiwan sud	Oryza	sativa	na	1	na	2	-	250	23.69781	120.960515	pays	-
TW0010	na	Asia	Taiwan	Nankang	Taipei	Oryza	sativa	na	1	na	2	-	161	25.0329694	121.5654177	ville	-
UR0002	1999	Americas	Uruguay	na	na	Oryza	sativa	El Paso 144	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0004	2002	Americas	Uruguay	na	na	Oryza	sativa	Tacuari	2	na	3	-	154	-32.522779	-55.765835	pays	-
UR0005	1998	Americas	Uruguay	na	UEPL	Oryza	sativa	CT13063 CA98	na	na	2	-	192	-33.2809537	-54.1843929	ville	-
UR0006	1998	Americas	Uruguay	na	Rio Branco	Oryza	sativa	Tacuarin 6-1	na	na	2	-	192	-32.5931047	-53.3744131	province	-
UR0007	1997	Americas	Uruguay	na	na	Oryza	sativa	Blue Belle	na	na	2	-	192	-32.522779	-55.765835	pays	-
UR0008	1997	Americas	Uruguay	na	UEPL-Prog Arroz	Oryza	sativa	SCT10885	na	na	2	-	192	-33.2809537	-54.1843929	ville	-
UR0009	1999	Americas	Uruguay	na	na	Oryza	sativa	L3019	na	na	2	-	192	-32.522779	-55.765835	pays	-
UR0010	2003	Americas	Uruguay	na	na	Oryza	sativa	Tacuari	1	na	2	-	192	-32.522779	-55.765835	pays	-
UR0011	2003	Americas	Uruguay	na	na	Oryza	sativa	Tacuari	1	na	2	-	192	-32.522779	-55.765835	pays	-
UR0012	1999	Americas	Uruguay	na	na	Oryza	sativa	L3194	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0013	2003	Americas	Uruguay	na	na	Oryza	sativa	Tacuari	na	na	2	-	192	-32.522779	-55.765835	pays	-
UR0014	1999	Americas	Uruguay	na	na	Oryza	sativa	L2887	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0015	1999	Americas	Uruguay	na	na	Oryza	sativa	L2995	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0016	1999	Americas	Uruguay	na	na	Oryza	sativa	L3128	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0017	2001	Americas	Uruguay	na	Cebollati	Oryza	sativa	Tacuari	na	na	2	-	192	-32.522779	-55.765835	pays	-
UR0018	2001	Americas	Uruguay	na	Cebollati	Oryza	sativa	Tacuarí	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0019	2001	Americas	Uruguay	na	na	Oryza	sativa	El Paso 144	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0020	2001	Americas	Uruguay	na	na	Oryza	sativa	Tacuari	na	na	3	-	154	-32.522779	-55.765835	pays	-
UR0021	2002	Americas	Uruguay	na	na	Oryza	sativa	L1882	na	na	2	-	192	-32.522779	-55.765835	pays	-
US0001	1987	Americas	USA	Arkansas	Stuttgart	Oryza	sativa	Mars	na	non-female	1	Int	2	34.5003748	-91.5526281	ville	-
US0002	1987	Americas	USA	Arkansas	Pine Tree	Oryza	sativa	Mars	na	non-female	1	Int	2	33.005046	-92.557518	pays	-
US0003	1987	Americas	USA	Arkansas	Stuttgart	Oryza	sativa	Newbonnet	na	non-female	1	Int	2	34.5003748	-91.5526281	ville	-
US0004	1987	Americas	USA	Arkansas	Pine Tree	Oryza	sativa	Usen	na	non-female	1	Int	2	33.005046	-92.557518	pays	-
US0005	1987	Americas	USA	Arkansas	Cross Cnty	Oryza	sativa	Tebonnet	1	na	2	-	13	35.2998913	-90.8294002	province	-
US0006	1988	Americas	USA	Mississipi	Stoneville	Oryza	sativa	M201	na	non-female	1	Int	2	33.4240055	-90.9151014	ville	-
US0008	1988	Americas	USA	Texas	Beaumont	Oryza	sativa	Dawn	na	na	2	-	73	30.080174	-94.1265562	ville	-
US0009	1988	Americas	USA	Texas	Warton Cnty	Oryza	sativa	Gulfmont	1	na	2	-	73	29.3690932	-96.1526985	ville	-
US0010	1988	Americas	USA	Texas	Jackson Cnty	Oryza	sativa	Lemont	1	na	2	-	73	28.9340886	-96.5356449	province	-
US0030	1975	Americas	USA	Louisiana	na	Oryza	sativa	na	2	na	3	-	72	30.9842977	-91.9623327	province	-
US0032	1964	Americas	USA	Louisiana	na	Oryza	sativa	na	na	non-female	4	-	158	30.9842977	-91.9623327	province	yes
US0033	1975	Americas	USA	Texas	na	Oryza	sativa	na	na	na	2	-	73	31.9685988	-99.9018131	province	-
US0035	1974	Americas	USA	Texas	na	Oryza	sativa	na	1	na	2	-	231	31.9685988	-99.9018131	province	-
US0036	1974	Americas	USA	Louisiana	na	Oryza	sativa	na	1	na	2	-	19	30.9842977	-91.9623327	province	-
US0039	1985	Americas	USA	Louisiana	na	Oryza	sativa	na	na	na	2	-	150	30.9842977	-91.9623327	province	-
US0041	1985	Americas	USA	Mississipi	na	Oryza	sativa	na	2	non-female	1	Int	98	32.3546679	-89.3985283	province	yes
US0052	1959	Americas	USA	Louisiana	Calcasieu	Oryza	sativa	Zenith	na	na	4	-	158	30.2089286	-93.3388917	province	-
US0089	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	2	na	2	-	21	34.7839795	-91.8998611	ville	-
US0090	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	1	non-female	1	Int	2	34.7839795	-91.8998611	ville	yes
US0091	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	2	na	2	-	229	34.7839795	-91.8998611	ville	-
US0098	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	1	non-female	2	-	224	34.7839795	-91.8998611	ville	yes
US0099	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	2	non-female	1	Int	2	34.7839795	-91.8998611	ville	-
US0100	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	2	non-female	1	Int	2	34.7839795	-91.8998611	ville	-
US0106	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	1	na	2	-	13	34.7839795	-91.8998611	ville	-
US0107	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	2	non-female	1	Int	98	34.7839795	-91.8998611	ville	-
US0119	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	Newbonnet	na	na	2	-	13	34.7839795	-91.8998611	ville	-
US0160	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	V.Mars	1	na	2	-	116	34.7839795	-91.8998611	ville	-
US0164	1992	Americas	USA	Arkansas	Lonoke	Oryza	sativa	V.Mars	1	na	2	-	116	34.7839795	-91.8998611	ville	-
VN0002	2006	Americas	Venezuela	Guarico	Carretera nacional Calabozo	Oryza	sativa	D-Sativa	na	na	3	-	169	8.8985763	-67.4477399	ville	-
VT0001	1989	Asia	Vietnam	na	Hochi Minh	Oryza	sativa	na	2	na	3	-	200	10.8230989	106.6296638	ville	-
VT0003	1994	Asia	Vietnam	Thuanthien Hue	Huong Thuy Thuy Tan	Oryza	sativa	Nep Lun , Sticky rice	2	non-female	1	Yule	247	16.4651937	107.6416471	ville	yes
VT0005	1994	Asia	Vietnam	Thuanthien Hue	Huong Thuy Thuy Tan	Oryza	sativa	IR 38	1	na	3	-	63	16.4651937	107.6416471	ville	-
VT0009	1994	Asia	Vietnam	Thuanthien Hue	Huong Viny Thuy Duong	Oryza	sativa	IR17494	2	na	3	-	63	16.4411867	107.620996	ville	-

VT0010	1994	Asia	Vietnam	Hue	Thuy Phuong	Oryza	sativa	Sticky rice	2	na	1	Yule	247	16.4782764	107.58291	ville	-
VT0011	2001	Asia	Vietnam	Ninh binh	Hoa lu	Oryza	sativa	na	na	na	3	-	154	20.248813	105.9112424	ville	-
VT0014	2001	Asia	Vietnam	Ninh binh	Hoa lu	Oryza	sativa	na	2	na	3	-	182	20.248813	105.9112424	ville	-
VT0017	2001	Asia	Vietnam	Thai nguyen	Pho yen	Oryza	sativa	DT13	2	na	3	-	81	21.413875	105.8486183	ville	-
VT0027	2003	Asia	Vietnam	Phu Ly	na	Oryza	sativa	Sticky rice	2	non-female	1	Laos	26	20.5316929	105.9176291	ville	yes
VT0030	2003	Asia	Vietnam	Hanam	Kim Bang	Oryza	sativa	Sticky rice	1	non-female	1	Laos	74	20.5633623	105.8562473	province	yes

DISCUSSION GENERALE

L'étude de la structure génétique de l'espèce fongique phytopathogène *Pyricularia oryzae* réalisée au cours de cette thèse a permis la caractérisation de certaines lignées qui composent cette espèce mais également de mieux comprendre l'évolution et la propagation de cet agent pathogène menaçant la sécurité alimentaire mondiale. Nos connaissances sur la structure des lignées hôtes-spécifiques ont permis de développer des tests de diagnostic moléculaire essentiels pour éviter l'extension à de nouvelles aires géographiques de la pyriculariose du blé. La structure existant au sein de la lignée Oryza a, quant à elle, été affinée et décrite à l'échelle globale permettant l'étude de la propagation mondiale de l'agent pathogène sur riz, de son évolution ainsi que des facteurs impliqués dans la structuration génétique de cette lignée.

1- Intégration de la détection infraspécifique de la lignée Triticum dans les normes de biosécurité actuelles

Aujourd'hui les mesures de biosécurité sont basées sur l'identification de taxons, souvent des espèces, préalablement décrits et nommés (McTaggart et al., 2016). Toutefois, définir avec précision les limites des taxons ciblés n'est souvent pas trivial chez les champignons (McTaggart et al., 2016; Steenkamp et al., 2018). Entre espèces fongiques, des échanges de matériels génétiques par transferts horizontaux ont souvent pu être décrits (Barreiro et al., 2019; Fitzpatrick, 2012; Soanes and Richards, 2014) ainsi que de multiples phénomènes d'hybridations interspécifiques (Leducq et al., 2016; Shoji et al., 2015; Stukenbrock et al., 2012). Les limites entre espèces fongiques sont alors décrites comme semi-perméables (Steenkamp et al., 2018). Cette perméabilité semble être d'autant plus grande que les espèces sont apparentées. En effet, comme mentionné en introduction, la spéciation est un processus continu au long duquel les différences génétiques et phénotypiques entre lignées s'accumulent formant une barrière aux flux de gènes de plus en plus imperméable. Ainsi, des échanges de matériels génétiques sont plus probables entre espèces fongiques génétiquement proches qu'entre espèces génétiquement éloignées (Steenkamp et al., 2018). Dans le cas de taxons identifiés au sein même d'une espèce, comme c'est le cas des lignées hôte-spécifiques existant au sein de *P. oryzae,* les échanges de matériel génétique entre taxons n'en sont que plus probables. En effet, des flux de gènes ont pu être observés entre les lignées hôte-spécifiques de *P. oryzae* (Gladieux et al., 2018b) et une possible reproduction sexuée entre individus de différentes lignées a pu être démontré par des croisements en condition de laboratoire (Zeigler, 1998). Dans la perspective de mise en place de mesures afin d'éviter l'introduction d'une maladie dévastatrice dans des régions qui en sont actuellement exemptes, il existe alors un fort décalage entre le fait de savoir que les limites entre ces

lignées demeureront floues et la nécessité de disposer d'outils « infra-spécifiques » permettant la détection d'une de ces lignées au sein de l'espèce *P. oryzae*.

Au cours de cette thèse j'ai utilisé une approche de comparaison de génomes afin d'identifier des polymorphismes spécifiques de la lignée Triticum de *P. oryzae*, responsable des épidémies de pyriculariose du blé. Plusieurs tests de détection ciblant plusieurs de ces polymorphismes ont ainsi pu être développés (C17, C45, LAMP n°5) et sont venus s'ajouter aux tests existants ciblant le locus MoT3 (Pieck et al. 2017; Yasuhara-Bell et al. 2019). Le développement de ces tests a permis d'améliorer la détection des isolats responsables de la pyriculariose du blé. L'ensemble des tests (C17, C45 et LAMP n°5) a en effet montré une parfaite inclusivité sur les isolats testés (détection de l'ensemble des isolats responsables des épidémies de pyriculariose du blé) ainsi qu'une forte spécificité. Toutefois, individuellement ces tests ont également montré leurs limites. Les loci ciblés par les tests C17 et LAMP n°5 sont présents dans certains isolats n'appartenant pas à la lignée Triticum entrainant des faux-positifs lors de leur utilisation. Ceci peut être dû à un flux de gènes entre les lignées hôte-spécifiques ou à un tri incomplet des lignées. Le polymorphisme ciblé par le test C45 quant à lui, ne consiste qu'en un seul SNP ce qui peut entrainer de faibles hybridations des amorces et des amplifications tardives des ADN non-cibles pouvant faussement être interprétées comme un signal positif. La complémentarité de ces tests, ciblant de multiples loci, semble toutefois permettre une détection complètement inclusive et spécifique de la lignée Triticum sur l'ensemble des isolats testés jusqu'à présent. L'approche consistant à cibler de multiples loci pourrait donc être une méthode efficace pour réaliser des diagnostics dans des cas de barrières fortement perméables entre taxons comme celles observées dans notre cas.

Une autre approche pourrait être de cibler directement les allèles de gènes conférant aux isolats de *P. oryzae* la pathogénicité sur blé. De façon générale, la stratégie consistant à cibler les régions qui confèrent la pathogénicité sur une plante hôte dépend très fortement de l'histoire évolutive des régions en question, et en particulier de leur mode de transmission, et nécessite au préalable une étude approfondie de l'évolution de ces marqueurs au sein et entre les lignées. Les gènes *Pwt3* et *Pwt4* sont importants pour la pathogénicité sur blé car codent pour des protéines d'avirulence reconnues par les protéines de résistance Rwt3 et Rwt4, respectivement, présentes dans certaines variétés de blé (Inoue et al., 2017). Inoue et al. (2017) ont d'ailleurs montré que les populations de *P. oryzae* infectant le blé ont évolué au cours du temps vers une perte fonctionnelle du gène *Pwt3* par différents évènements d'insertion de transposons ou de mutations ponctuelles dans la séquence du gène. Le gène *Pwt4* est quant à lui souvent absent des isolats de *P. oryzae* (Inoue et al., 2017). Dans le cadre de cette thèse nous avons étudié les allèles du gène *Pwt3* présents dans les génomes des isolats appartenant aux différentes lignées hôtes spécifiques, et avons mis en évidence de multiples allèles

avec des profils complexes au sein de la lignée *Triticum* (insertions de larges séquences au sein du gène, séquences du gène fragmentée en deux ou trois fragments retrouvés sur différents scaffolds des génomes assemblés) compatibles avec une forte pression de sélection favorisant les allèles mutés du gène. De plus, tout comme cela avait été mis en évidence par Inoue et al. (2017), nous avons identifié des allèles partagés entre des isolats de la lignée *Triticum* et des isolats d'autres lignées hôte spécifique, dont l'allèle B virulent (Inoue et al., 2017) porté par l'isolat SSFL02 appartenant à la lignée Stenotaphrum (données non présentées). La multitude des allèles de *Pwt3* au sein de la lignée Triticum ainsi que la présence de certains de ces allèles dans d'autres lignées hôte spécifiques (Inoue et al., 2017; Thierry et al., 2019), ne permet donc malheureusement pas de cibler la séquence de ce gène pour le développement d'un test de détection de la lignée Triticum. Sans remettre en cause le scénario évolutif proposé pour *Pwt3* dans les populations de *P. oryzae* infectant le blé, ni son importance pour l'infection des variétés de blé portant le gène de résistance correspondant, il est difficile de penser que l'évolution de ce gène soit seule responsable de l'émergence de la pyriculariose du blé. Si tel était le cas, les isolats des autres lignées hôtes spécifiques portant des allèles virulents du gène auraient pu être responsables d'épidémies sur blé. De plus, la multitude des allèles au sein de la lignée Triticum et leur généalogie ne semblent pas compatibles avec un unique saut d'hôte qui aurait été permis par une mutation de ce gène. Il est toutefois intéressant de noter que les polymorphismes spécifiques de la lignée Triticum qui ont été identifiés par notre approche de comparaison de génomes sont localisés dans un nombre très restreint de régions du génome. Le fait que toutes ces positions soient regroupées pourrait être dû à l'existence d'un déséquilibre de liaison entre ces régions et un locus sous sélection positive dans la lignée Triticum. D'autres gènes importants pour la pathogénicité sur blé et directement impliqués dans le saut d'hôte, pourraient être présents dans ces régions.

L'approche de comparaison de génomes utilisée dans cette thèse pour la recherche de multiples loci spécifiques d'une lignée au sein même d'une espèce a démontré tout son intérêt. Le développement des techniques de séquençage et la baisse de leur coût a drastiquement augmenté la quantité de données génomiques disponibles pour les espèces fongiques. Ainsi, des méthodes similaires pour le développement de tests de diagnostic à un niveau infraspécifique dans le cas d'autres espèces fongiques pourraient être utilisées.

2- Que sait-on de la propagation de *P. oryzae* à l'échelle mondiale ?

Le transport de *P. oryzae* à l'échelle globale est causé par le transport de matériel biologique contaminé, principalement de semences (Faivre-Rampant et al., 2013; Goulart, 1988). En effet sans l'action de l'Homme les spores du champignon ne sont capables de se propager qu'à faible distance (Tharreau et al., 2009). Ainsi, la pyriculariose du blé a par exemple été introduite pour la première fois en Asie en 2016 certainement par une importation de semences contaminée depuis l'Amérique du sud

où a émergé cette maladie (Islam et al., 2016). De même, des isolats de *P. oryzae* capables d'infecter les variétés de riz portant le gène de résistance *Pi33* se sont très certainement propagés mondialement à partir d'un nombre restreint de zones où le contournement de résistance a eu lieu (Tharreau et al., 2009). Dans cette thèse l'étude de près de 900 isolats prélevés dans plus de 51 pays et génotypés pour plus de 5500 marqueurs nous permet de mettre en évidence l'ampleur de cette propagation intercontinentale concernant la lignée Oryza. Des isolats ayant des génotypes identiques (pour l'ensemble des marqueurs étudiés), et donc fortement apparentés, ont pu être identifiés dans de multiples pays impliquant un transport de l'agent pathogène d'un pays à l'autre. Ces mouvements de l'agent pathogène entre pays semblent plus importants à un niveau intracontinental en Amérique, en Europe ainsi qu'en Asie mais de multiples échanges intercontinentaux sont également détectés. Notre étude a également mis en évidence un cas intéressant, celui du continent Européen. Les études précédentes avaient montré la présence d'une unique lignée clonale (à une exception près) qui avait conduit à formuler l'hypothèse d'une unique introduction de l'agent pathogène sur le continent et de flux très limités de *P. oryzae* vers ce continent. Notre étude a montré que les génotypes des isolats européens sont distribués sur l'ensemble de la phylogénie de la lignée 2, présents sur des branches de la phylogénie comportant également des isolats non européens, et que des génotypes identiques sont partagés entre l'Europe et d'autres continents suggérant un grand nombre d'échanges entre l'Europe et le reste du monde. Ce profil pourrait s'expliquer par trois scénarios :

1) Une unique introduction en Europe de la lignée, une diversification de l'agent pathogène au sein du continent suivie de multiples exportations à travers le monde. Ceci impliquerait un transport à sens unique de l'agent pathogène ;
2) Plusieurs introductions ont eu lieu mais seulement d'isolats appartenant à la lignée 2 (possible par exemple si les variétés importées sont fortement associées à la lignée 2) ;
3) Des introductions multiples de différentes lignées ont eu lieu mais l'existence de pressions de sélections fortes empêche la survie de lignées autre que la lignée 2 sur le territoire européen.

Différentes expérimentations pourraient être mises en place pour évaluer ces hypothèses : l'étude du matériel végétal actuellement importé et exporté d'Europe (par exemple par l'évaluation du niveau de contamination des semences de riz, l'évaluation des lignées associées à ces semences) permettraient de faire le point sur les échanges ayant actuellement lieu ; des études d'évolutions expérimentales dans des conditions mimant le climat et les écosystèmes européens pourraient permettre de déterminer la capacité de survie de chacune des lignées dans ces conditions.

Il est important de comprendre et d'éviter ces transports de l'agent pathogène sur de longues distances qui peuvent entrainer l'importation de nouveaux génotypes, plus agressifs ou infectants de nouveaux hôtes, et qui ne semblent pas être anecdotiques. Toutefois, dans l'analyse que nous avons

menée il n'y a pas d'échelle de temps. Il serait intéressant de déterminer la période et le laps de temps dans lesquels ont pu avoir lieu ces évènements de propagation. Des méthodes de datation moléculaires existent (Rieux and Balloux, 2016) et permettent par l'étude de la divergence génétique entre les séquences d'obtenir des inférences sur les temps de divergence, les taux de substitution, la démographie passée ou l'âge de certaines mutations à diverses échelles spatio-temporelles.

D'autre part, si la propagation des isolats appartenant à la lignée Oryza ou à la lignée Triticum a été étudiée, la situation sur d'autres espèces hôtes reste mal connue. Certaines propagations ont pourtant déjà eu un impact important tel que l'apparition des épidémies de pyriculariose sur Lolium en France en 2018 faisant des dégâts sur le ray-grass cultivé pour des gazons récréatifs (Milazzo et al., 2019). De plus certains isolats de la lignée Lolium étant capables d'infection opportunistes sur des variétés de blé (Farman et al., 2016; Inoue et al., 2017), un nouveau saut d'hôte de ces isolats vers le blé est un risque inquiétant pour l'Europe.

3- Association entre propagation hors du centre d'origine et clonalité au sein de la lignée Oryza

P. oryzae est capable de se reproduire sexuellement ou asexuellement. La reproduction sexuée n'a jamais été observée dans la nature. Cependant, les preuves indirectes de reproduction sexuée (caractéristiques génétiques et biologiques des populations de la lignée Oryza de *P. oryzae*) ne sont rencontrées qu'au centre présumé d'origine de l'agent pathogène et ne concerne que la lignée 1 recombinante (cette étude ; Gladieux et al., 2018b; Saleh et al., 2012). Les études sur les populations mondiales montrent que les lignées clonales 2, 3 et 4 dominent en dehors du centre d'origine de l'agent pathogène (cette étude ; Saleh et al. 2014 ; Tharreau et al. 2009). Ceci est cohérent avec de nombreuses études de populations de la lignée Oryza réalisées à des échelles locales et observant des populations clonales ayant un seul type sexuel dans un grand nombre de régions en dehors de l'Asie du Sud-Est (Chen et al., 1995; Consolo et al., 2005; Hemmati et al., 2005; Levy et al., 1993; Pagliaccia et al., 2018; Tharreau et al., 2009). Cependant, 10 génotypes (soit 40 isolats) appartenant à la lignée 1 ont été échantillonnés en dehors d'Asie. Fait intéressant, parmi ces 10 génotypes, trois ont été échantillonnés sur plusieurs continents et deux dans plusieurs pays du même continent, ce qui est compatible avec une propagation clonale de ces génotypes particuliers. La fertilité a été évaluée pour 24 de ces 40 isolats. Parmi eux, un seul isolat (GY0011) s'est révélé femelle-fertile, deux étaient mâle-fertiles mais femelle-stériles et enfin 21 étaient complètement stériles. Ces résultats montrent que, même en considérant les isolats appartenant à la lignée 1 recombinante, une propagation hors du centre d'origine de l'agent pathogène est fortement corrélée à une perte de fertilité. Plusieurs hypothèses pourraient expliquer la restriction de la reproduction sexuée à cette zone géographique.

La reproduction sexuée pourrait procurer un avantage sélectif dans cette région du monde où des variétés traditionnelles sont encore fortement utilisées entrainant une forte diversité génétique de l'hôte. D'autres conditions environnementales (conditions climatiques, systèmes de culture ...) pourraient également sélectionner ce mode de reproduction si celui-ci permet une meilleure survie de l'agent pathogène. Inversement la reproduction asexuée pourrait être sélectionnée en dehors du centre d'origine du champignon. La reproduction asexuée peut en effet procurer plusieurs avantages pour l'invasion de nouveaux environnements tels que (i) la possibilité de se multiplier sans besoin d'un partenaire compatible ; (ii) la production d'un grand nombre de spores permettant l'invasion du nouvel environnement principalement si le nouvel environnement ne présente pas de diversité ; (iii) éviter un fardeau génétique (« genetic load ») à savoir la recombinaison avec la population ancestrale non adaptée au nouvel environnement (Gladieux et al., 2015; Tibayrenc and Ayala, 2012). Aucune de ces hypothèses n'a toutefois été testée à ce jour.

4- Facteurs structurant les populations au sein de la lignée Oryza

A l'échelle globale, une forte corrélation entre conditions climatiques et aire de répartition des lignées est observée dans notre étude. Les variables climatiques étudiées concernent principalement des mesures de températures et d'humidité. Toutefois, des différences en termes de climat peuvent elle-même être corrélées à un grand nombre d'autres variables biotiques (espèces hôte cultivées, autres espèces pathogènes en compétition...) ou abiotiques (systèmes de cultures, UV…). Sur le continent européen par exemple les variétés de riz cultivées sont principalement de type Japonica tempéré car celles-ci sont adaptées aux conditions climatiques des zones tempérées. La corrélation observée entre lignées et climat peut donc être due à tous les facteurs de l'environnement.

La structuration par l'hôte des populations de *Pyricularia oryzae* infectant le riz (que ce soit les groupes génétiques de riz, les variétés ou la virulence vis-à-vis de gènes de résistance) a fait l'objet de nombreuses études (Chen et al., 1995, 2006; Don et al., 1999; Gladieux et al., 2018a; Goncalves et al., 2016; Javan-Nikkhah et al., 2004; Onaga et al., 2015; Park et al., 2008; Roumen et al., 1997; Thuan et al., 2006). Si une forte structuration par espèce-hôte, menant à la formation des lignées hôtes spécifiques, a pu être démontrée au sein de *P. oryzae,* l'impact de l'hôte sur la structure des populations au sein même de la lignée hôte spécifique Oryza semble moins marqué qu'attendu. Certaines études ont pu mettre en évidence une corrélation entre la structure génétique des populations et leur hôte. Ainsi certaines mettent en évidence un lien entre structure génétique et (i) virulence sur certaines variétés cultivées (Chen et al., 1995; Park et al., 2008), (ii) virulence sur des variétés différentielles (Roumen et al., 1997), ou (iii) virulence sur les groupes génétiques de riz Japonica et Indica (Gladieux et al., 2018a; Thuan et al., 2006) tandis qu'aucune corrélation visible n'a

pu être notée dans d'autres cas (Chen et al., 2006; Don et al., 1999; Onaga et al., 2015). Dans nos conditions expérimentales, il existe une claire différence de largeur de spectre d'hôtes en fonction des lignées étudiées. Les isolats appartenant à la lignée 2 montrent un spectre d'hôtes significativement réduit en n'infectant que les variétés hautement sensibles, confirmant les résultats précédemment obtenus (Gladieux et al., 2018a). Ces variétés sensibles appartiennent majoritairement au groupe génétique des riz Japonica tempérés. Ces mêmes variétés sont toutefois également sensibles à l'ensemble des autres lignées dans nos conditions expérimentales. Les spectres de virulence des lignées semblent donc se chevaucher. Le même type de résultats a été obtenu lors de précédentes études de l'adaptation des isolats de *P. oryzae* à leur hôte d'origine. Dans ces études, des isolats prélevés sur des riz Indica se sont révélés capables d'infecter à la fois des variétés Indica et des variétés Japonica tempérées sur lesquelles ils ont été réinoculées. Inversement, les isolats prélevés sur des riz Japonica ne se sont révélés virulents que sur les variétés Japonica (Gallet et al., 2016; Liao et al., 2016).

Les expérimentations visant à tester une adaptation à certaines conditions abiotiques telles que la température n'ont pas permis de mettre en évidence des profils d'adaptation tranchés et expliquant seuls la structure des populations. La croissance mycélienne des isolats semble suivre des tendances similaires pour toutes les lignées aux températures testées. On peut toutefois noter une croissance significativement plus basse des isolats de la lignées 4 par rapport aux autres lignées à 10°C et une croissance significativement plus rapide des isolats de la lignée 1 par rapport à toutes les autres lignées à toutes les autres températures testées (15°C, 20°C, 25°C et 30°C). Concernant la sporulation, la forte variabilité des mesures entre répétitions et entre isolats n'a pas permis de mettre en évidence des différences significatives entre lignées pour la plupart des températures étudiées. Cette évaluation de l'adaptation à la température à toutefois été faite en conditions artificielles très éloignées des conditions naturelles de vie du champignon. Parmi les différences majeures entre conditions naturelles et artificielles, on peut citer le fait qu'en conditions artificielles le champignon est en condition de culture optimale (milieu de culture riche, pas de stress liés à l'environnement) ou que plusieurs phases du cycle de vie du champignon ne sont pas représentées (phase de fixation des spores sur la plante hôte, phase d'infection de la plante hôte). Une adaptation à la température pourrait donc exister en condition naturelles sans être observables dans nos conditions expérimentales. L'humidité est un facteur abiotique qui n'a pas du tout été étudié jusqu'alors. Pourtant certaines études ont montré que le taux d'humidité est un facteur crucial pour le succès de l'infection de *P. oryzae* (30 minutes d'hydratation essentielles pour l'adhésion des spores sur les feuilles de la plante hôte) et pourrait entrainer une forte pression de sélection sur les populations (Ikeda et al., 2019).

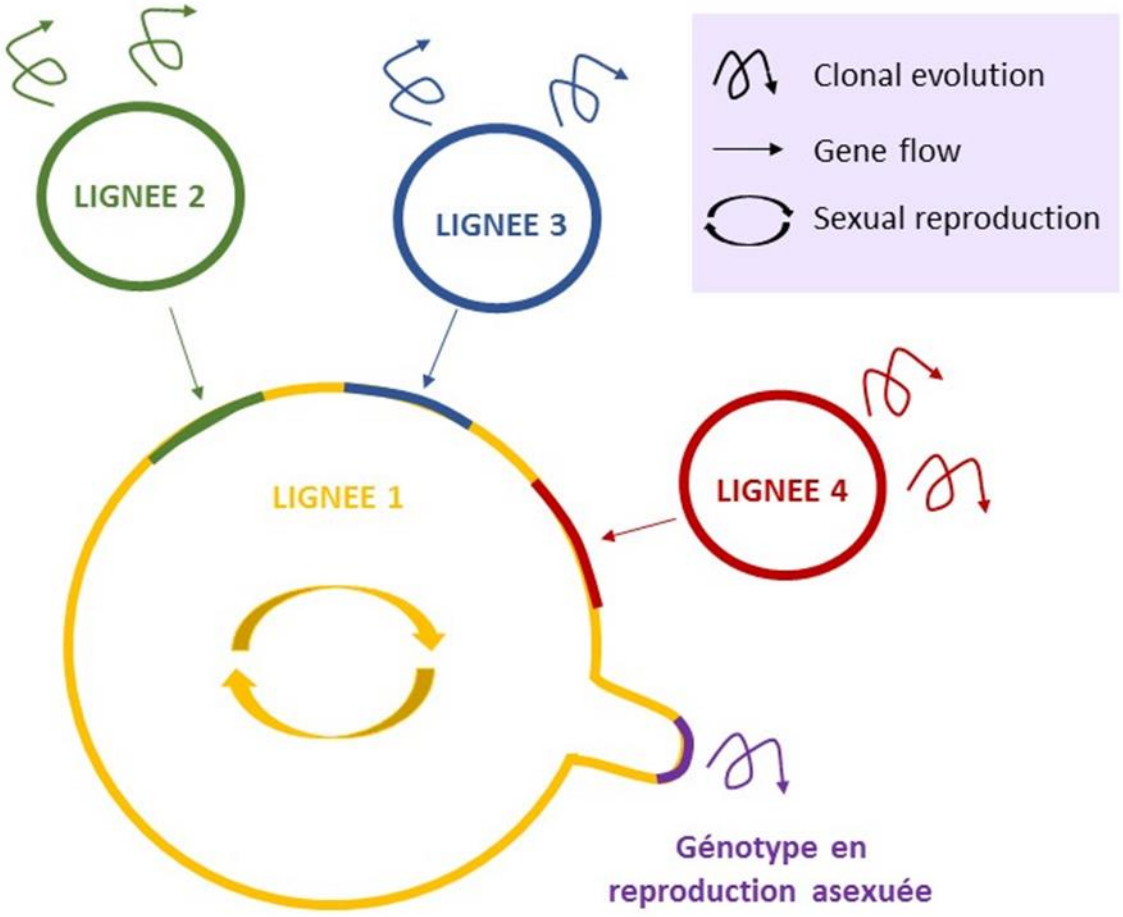


Figure 6: Schématisation de l'apparition, l'interaction et l'évolution des lignées majeures au sein de *P. oryzae*

La forte corrélation entre climats et aire de répartition des lignées indique certainement une adaptation des lignées à certaines conditions environnementales. Au cours de cette thèse nous avons recherché des adaptations différentielles des lignées à plusieurs facteurs environnementaux tels que l'hôte ou la température. Ces facteurs ont été étudiés de façons indépendantes les uns des autres mais il est fort probable que les différences d'aire de répartitions des lignées soient la résultante d'une somme d'adaptations à de multiples facteurs environnementaux. Des analyses multifactorielles et/ou des expérimentations en champ pourraient permettre de mieux comprendre les adaptations différentielles des lignées aux conditions environnementales.

Cette thèse montre également que le maintien de la structure génétique observée au sein de la lignée Oryza est dû aux fortes barrières à la reproduction sexuée existantes entre les lignées qui la composent. Nous avons en effet pu mettre en évidence des incompatibilités de types sexuels entre certaines lignées clonales, une perte de fertilité femelle généralisée au sein des lignées clonales ainsi que des incompatibilités post-appariement fortes entre lignées. L'ensemble de ces barrières semble empêcher tout flux de gènes entre lignées clonales et le réduire fortement entre lignées clonales et lignée recombinante.

5- Modèle d'évolution et des flux de gènes au sein de la lignée Oryza

La combinaison d'informations sur la structure des populations, les flux de gènes et les modes de reproduction nous permet de proposer un modèle d'interaction et d'évolution des lignées majeures existant au sein de la lignée Oryza (Figure 6). Les lignées 2, 3 et 4 sont hautement différenciées et fortement clonales. Aucune recombinaison ne se produit au sein de ces lignées ou entre ces lignées et les isolats évoluent indépendamment les uns des autres. La théorie prédit que de forts évènements de sélection façonnent de manière significative les populations clonales. Ainsi, les génotypes ayant la meilleure valeur sélective auront tendance à complètement remplacer la population initiale (Ali et al. 2010). Ce type d'évolution a pu être décrit aux États-Unis où elle semble avoir été causée par le déploiement de gènes de résistance dans de nouvelles variétés de riz (Wang et al. 2017). D'autre part, nous avons révélé des croisements possibles entre les lignées clonales et la lignée 1. Ces croisements semblent tout de même moins probables que les croisements au sein de la lignée 1. Ces flux de gènes réduits pourraient expliquer le profil génétique de la lignée 1 dont les isolats semblent porter certains marqueurs spécifiques des lignées clonales, ces marqueurs restant toutefois minoritaires. La sous-structure géographique observée dans la lignée 1 est compatible d'une part avec la reproduction sexuée au sein de cette lignée et d'autre part avec la dispersion des spores à courte distance conduisant à un isolement par la distance des populations (Tharreau et al., 2009). L'un des avantages

prédit de la reproduction sexuée sur le plan de l'évolution est la production de combinaisons géniques avantageuses (de Meeûs et al., 2007; Samils et al. 2001). Onze des 12 isolats les plus multivirulents lors des inoculations effectuées au cours de cette thèse (symptômes> 2 sur plus de 40 variétés testées) appartiennent effectivement à lignée 1. Par ailleurs, certains isolats présentant des génotypes strictement identiques, assignés à la lignée 1, et présentant une fertilité réduite ont pu être identifiés dans plusieurs pays différents. Ceci implique une propagation clonale de ces isolats de la lignée 1 à travers le monde. Ces génotypes particuliers devraient être caractérisés en utilisant davantage de marqueurs génétiques afin de quantifier plus précisément leur divergence par rapport aux autres populations de la lignée 1. Cependant, nous pouvons certainement déjà conclure que ces génotypes particuliers représentent des lignées clonales ayant récemment divergé. L'évolution, la propagation et les caractéristiques biologiques de ces lignées devrait être plus attentivement étudiées. L'apparition des lignées clonales 2, 3 et 4 pourrait avoir résulté d'un mécanisme similaire, à savoir, l'apparition au sein de la lignée recombinante de génotypes ayant une forte valeur sélective permettant la propagation des isolats en dehors du centre d'origine de l'agent pathogène par une reproduction asexuée associée à une perte de fertilité des isolats. L'étude des génotypes ayant récemment émergé de la population recombinante et qui commencent une propagation asexuée, ainsi que des études d'évolution expérimentales pourraient renseigner sur les phases précoces de l'émergence des lignées clonales et les facteurs entrainant cette émergence.

Références bibliographiques

Afouda, L., Wolf, G., and Wydra, K. (2009). Development of a Sensitive Serological Method for Specific Detection of Latent Infection of *Macrophomina phaseolina* in Cowpea. p.

Alberione, E. (2008). Evaluación de cultivares de trigo en siembra de verano. Campañas 2006/2007 - 2007/2008 | Instituto Nacional de Tecnología Agropecuaria (Santa Rosa, La Pampa).

Amato, A., Kooistra, W.H.C.F., Ghiron, J.H.L., Mann, D.G., Pröschold, T., and Montresor, M. (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. Protist *158*, 193–207.

Anagnostakis, S.L. (1982). Genetic Analyses of *Endothia parasitica*: Linkage Data for Four Single Genes and Three Vegetative Compatibility Types. Genetics *102*, 25–28.

Anderson, J.B., and Ullrich, R.C. (1979). Biological Species of *Armillaria mellea* in North America. Mycologia *71*, 402–414.

Andersson, L. (1990). The Driving Force: Species Concepts and Ecology. Taxon *39*, 375–382.

Avise, J.C., and Ball, J.R.M. (1990). Principles of genealogical concordance in species concepts and biological taxonomy. p.

Barea, G., and Toledo, J. (1996). Identificación y zonificación de piricularia o bruzone (*Pyricularia oryzae*) en el cultivo del trigo en el dpto. de Santa Cruz. CIAT Inf. Téc. Proy. Investig. Trigo 76–86.

Barreiro, C., Gutiérrez, S., and Olivera, E.R. (2019). Fungal Horizontal Gene Transfer: A History Beyond the Phylogenetic Kingdoms. In Horizontal Gene Transfer: Breaking Borders Between Living Kingdoms, T.G. Villa, and M. Viñas, eds. (Cham: Springer International Publishing), pp. 315–336.

Baum, D., and Shaw, K.L. (1995). Genealogical perspectives on the species problem. p.

Bennett, R., and Turgeon, G. (2017). Fungal Sex: The Ascomycota. In The Fungal Kingdom, Heitman, Howlett, Crous, Stukenbrock, James, and Gow, eds. (American Society of Microbiology), pp. 117–145.

Bernet, J. (1965). Mode d'action des gènes de "barrage" et relation entre l'incompatibilité cellulaire et l'incompatibilité sexuelle chez "*Podospora anserina*." p.

Besl, H., and Bresinsky, A. (1997). Chemosystematics of Suillaceae and Gomphidiaceae (suborder Suillineae). Plant Syst. Evol. *206*, 223–242.

Billiard, S., López-Villavicencio, M., Devier, B., Hood, M.E., Fairhead, C., and Giraud, T. (2011). Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. Biol. Rev. Camb. Philos. Soc. *86*, 421–442.

Broeders, S., Huber, I., Grohmann, L., Berben, G., Taverniers, I., Mazzara, M., Roosens, N., and Morisset, D. (2014). Guidelines for validation of qualitative real-time PCR methods. Trends Food Sci. Technol. *37*, 115–126.

Burger, W.C. (1975). The Species Concept in Quercus. Taxon *24*, 45–50.

Casselton, L.A., and Olesnicky, N.S. (1998). Molecular genetics of mating recognition in basidiomycete fungi. Microbiol. Mol. Biol. Rev. MMBR *62*, 55–70.

Chen, D., Zeigler, R., Leung, H., and Nelson, R. (1995). Population Structure of *Pyricularia grisea* at 2 Screening Sites in the Philippines. Phytopathology *85*, 1011–1020.

Chen, E.C.H., Mathieu, S., Hoffrichter, A., Sedzielewska-Toro, K., Peart, M., Pelin, A., Ndikumana, S., Ropars, J., Dreissig, S., Fuchs, J., et al. (2018). Single nucleus sequencing reveals evidence of inter-nucleus recombination in arbuscular mycorrhizal fungi. Elife *7*, e39813.

Chen, Q.H., Wang, Y.C., and Zheng, X.B. (2006). Genetic Diversity of Magnaporthe grisea in China as Revealed by DNA Fingerprint Haplotypes and Pathotypes. J. Phytopathol. *154*, 361–369.

Chiocchetti, A., Sciaudone, L., Durando, F., Garibaldi, A., and Migheli, Q. (2001). PCR Detection of *Fusarium oxysporum* f. sp. *basilici* on Basil. Plant Dis. *85*, 607–611.

Choi, J., Kim, H., and Lee, Y.-H. (2013). Comparative Analysis of the Korean Population of *Magnaporthe oryzae* by Multilocus Mierosatellite Typing. Plant Pathol. J. *29*, 435–439.

Coelho, M., Bakkeren, G., Sun, S., Hood, M., and Giraud, T. (2017). Fungal Sex: The Basidiomycota. In The Fungal Kingdom, Heitman, Howlett, Crous, Stukenbrock, James, and Gow, eds. (American Society of Microbiology), pp. 147–175.

Consolo, V.F., Cordo, C.A., and Salerno, G.L. (2005). Mating-type Distribution and Fertility Status in *Magnaporthe grisea* Populations from Argentina. Mycopathologia *160*, 285–290.

Correll, J.C., Harp, T.L., Guerber, J.C., Zeigler, R.S., Liu, B., Cartwright, R.D., and Lee, F.N. (2000). Characterization of *Pyricularia grisea* in the United States Using Independent Genetic and Molecular Markers. Phytopathology *90*, 1396–1404.

Couch, B.C., Fudal, I., Lebrun, M.-H., Tharreau, D., Valent, B., van Kim, P., Nottéghem, J.-L., and Kohn, L.M. (2005). Origins of Host-Specific Populations of the Blast Pathogen *Magnaporthe oryzae* in Crop Domestication With Subsequent Expansion of Pandemic Clones on Rice and Weeds of Rice. Genetics *170*, 613–630.

Coyne, J.A., and Orr, H.A. (1989). Patterns of Speciation in Drosophila. Evolution *43*, 362–381.

Coyne, J.A., and Orr, H.A. (2004). Speciation (Sinauer).

Cracraft, J. (1983). Species Concepts and Speciation Analysis. In Current Ornithology, R.F. Johnston, ed. (Boston, MA: Springer US), pp. 159–187.

Crous, P.W., Hawksworth, D.L., and Wingfield, M.J. (2015). Identifying and Naming Plant-Pathogenic Fungi: Past, Present, and Future. Annu. Rev. Phytopathol. *53*, 247–267.

Cruz, C.D., and Valent, B. (2017). Wheat blast disease: danger on the move. Trop. Plant Pathol. *42*, 210–222.

Darwin, C. (1859). On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life.

De Queiroz, K. (2007). Species Concepts and Species Delimitation. Syst. Biol. *56*, 879–886.

Dean, R., Van Kan, J. a. L., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol. *13*, 414–430.

Debuchy, R., Berteaux-Lecellier, V., and Silar, P. (2010). Mating Systems and Sexual Morphogenesis in Ascomycetes. In Cellular and Molecular Biology of Filamentous Fungi, (ASM Press), pp. 501–536.

De Meeûs, T., Prugnolle F., et Agnew P. (2007). Asexual Reproduction: Genetics and Evolutionary Aspects. Cell. Mol. Life Sci. *11,* 1355-72.

Dettman, J.R., Jacobson, D.J., and Taylor, J.W. (2003). A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. Evol. Int. J. Org. Evol. *57*, 2703–2720.

Dettman, J.R., Sirjusingh, C., Kohn, L.M., and Anderson, J.B. (2007). Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. Nature *447*, 585–588.

Dobzhansky, T. (1950). Mendelian Populations and Their Evolution. Am. Nat. *84*, 401–418.

Doehlemann, G., Ökmen, B., Zhu, W., and Sharon, A. (2017). Plant Pathogenic Fungi. Microbiol. Spectr. *5*.

Don, L.D., Tosa, Y., Nakayashiki, H., Mayama, S., and Kusaba, M. (1999). Population Structure of the Rice Blast fungus in Japa, examined by DNA fingerprinting. Ann Phytopathol Soc Jpn *65*, 15–24.

Donoghue, M.J. (1985). A Critique of the Biological Species Concept and Recommendations for a Phylogenetic Alternative. The Bryologist *88*, 172–181.

Drès, M., and Mallet, J. (2002). Host races in plant–feeding insects and their importance in sympatric speciation. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *357*, 471–492.

Faivre-Rampant, O., Geniès, L., Piffanelli, P., and Tharreau, D. (2013). Transmission of rice blast from seeds to adult plants in a non-systemic way. Plant Pathol. *62*, 879–887.

Farman, M., Peterson, G., Chen, L., Starnes, J., Valent, B., Bachi, P., Murdock, L., Hershman, D., Pedley, K., Fernandes, J.M., et al. (2017). The Lolium Pathotype of *Magnaporthe oryzae* Recovered from a Single Blasted Wheat Plant in the United States. Plant Dis. *101*, 684–692.

Fitzpatrick, D.A. (2012). Horizontal gene transfer in fungi. FEMS Microbiol. Lett. *329*, 1–8.

Fontaneto, D., and Barraclough, T.G. (2015). Do Species Exist in Asexuals? Theory and Evidence from *Bdelloid Rotifers*. Integr. Comp. Biol. *55*, 253–263.

Fontaneto, D., Herniou, E.A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., and Barraclough, T.G. (2007). Independently Evolving Species in Asexual *Bdelloid Rotifers*. PLOS Biol. *5*, e87.

Fournier, E., Giraud, T., Albertini, C., and Brygoo, Y. (2005). Partition of the *Botrytis cinerea* complex in France using multiple gene genealogies. Mycologia *97*, 1251–1267.

Friesen, T.L., Stukenbrock, E.H., Liu, Z., Meinhardt, S., Ling, H., Faris, J.D., Rasmussen, J.B., Solomon, P.S., McDonald, B.A., and Oliver, R.P. (2006). Emergence of a new disease as a result of interspecific virulence gene transfer. Nat. Genet. *38*, 953–956.

Frisvad, J.C., and Filtenborg, O. (1990). Secondary Metabolites as Consistent Criteria in *Penicillium* Taxonomy and a Synoptic Key to *Penicillium Subgenus Penicillium*. In Modern Concepts in *Penicillium* and *Aspergillus* Classification, R.A. Samson, and J.I. Pitt, eds. (Boston, MA: Springer US), pp. 373–384.

Funk, D.J., Nosil, P., and Etges, W.J. (2006). Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. Proc. Natl. Acad. Sci. *103*, 3209–3213.

Gallet, R., Fontaine, C., Bonnot, F., Milazzo, J., Tertois, C., Adreit, H., Ravigné, V., Fournier, E., and Tharreau, D. (2016). Evolution of Compatibility Range in the Rice− *Magnaporthe oryzae* System: An Uneven Distribution of R Genes Between Rice Subspecies. Phytopathology *106*, 348–354.

Garibaldi, A., Gilardi, G., and Gullino, M.L. (2004). Seed transmission of *Fusarium oxysporum f.sp.lactucae.* Phytoparasitica *32*, 61–65.

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic Structure and Diversity in *Oryza sativa* L. Genetics *169*, 1631–1638.

Gavrilets, S. (2004). Fitness Landscapes and the Origin of Species (MPB-41) (Princeton University Press).

Giraud, T. (2006). Speciation: Selection against migrant pathogens: the immigrant inviability barrier in pathogens.

Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D.M., and Hood, M.E. (2008a). Speciation in fungi. Fungal Genet. Biol. FG B *45*, 791–802.

Giraud, T., Yockteng, R., López-Villavicencio, M., Refrégier, G., and Hood, M.E. (2008b). Mating System of the Anther Smut Fungus *Microbotryum violaceum*: Selfing under Heterothallism. Eukaryot. Cell *7*, 765–775.

Giraud, T., Gladieux, P., and Gavrilets, S. (2010). Linking the emergence of fungal plant diseases with ecological speciation. Trends Ecol. Evol. *25*, 387–395.

Gladieux, P., Ravel, S., Rieux, A., Cros-Arteil, S., Adreit, H., Milazzo, J., Thierry, M., Fournier, E., Terauchi, R., and Tharreau, D. (2018a). Coexistence of Multiple Endemic and Pandemic Lineages of the Rice Blast Pathogen. MBio *9*, e01806-17.

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J.L.N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., et al. (2018b). Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the Rice Blast Fungus *Magnaporthe oryzae*. MBio *9*, e01219-17.

Glass, N. (2004). Hyphal homing, fusion and mycelial interconnectedness. Trends Microbiol. *12*, 135–141.

Goncalves, F.J., Corsi de Filippi, M.C., Silva Lobo, V.L., Araujo, L.G., Silva, G.B., Guedes Coelho, A.S., and Prabhu, A.S. (2016). Polymorphism Detection by Microsatellite Markers in a *Magnaporthe oryzae* Population From Different Geographical Areas of Brazil. J. Phytopathol. *164*, 620–630.

Goulart, A.C.P. (1988). Doenças transmitidas por sementes de trigo en Mato Grosso do Sul, e Testes de laboratório para identificação de *Pyricularia oryzae* Cav. en sementes de trigo.

Gupta, D.R., Avila, C.S.R., Win, J., Soanes, D.M., Ryder, L.S., Croll, D., Bhattacharjee, P., Hossain, Md.S., Mahmud, N.U., Mehebub, Md.S., et al. (2019). Cautionary Notes on Use of the MoT3 Diagnostic Assay for *Magnaporthe oryzae* Wheat and Rice Blast Isolates. Phytopathology *109*, 504–508.

Hahn, M.W., De Bie, T., Stajich, J.E., Nguyen, C., and Cristianini, N. (2005). Estimating the tempo and mode of gene family evolution from comparative genomic data. Genome Res. *15*, 1153–1160.

Hemmati, R., Javan-Nikkhah, M., Okovvat, S.M., and Ghazanfari, K. (2005). Study on genetic diversity of *Magnaporthe grisea* using PCR and determination of the mating type alleles distribution in Mazandaran province, Iran. Commun. Agric. Appl. Biol. Sci. *70*, 311–313.

Hofstatter, P.G., and Lahr, D.J.G. (2019). All Eukaryotes Are Sexual, unless Proven Otherwise: Many So-Called Asexuals Present Meiotic Machinery and Might Be Able to Have Sex. BioEssays *41*, 1800246.

Ikeda, K., Park, P., and Nakayashiki, H. (2019). Cell biology in phytopathogenic fungi during host infection: commonalities and differences. J. Gen. Plant Pathol. *85*, 163–173.

Inoue, Y., Vy, T.T.P., Yoshida, K., Asano, H., Mitsuoka, C., Asuke, S., Anh, V.L., Cumagun, C.J.R., Chuma, I., Terauchi, R., et al. (2017). Evolution of the wheat blast fungus through functional losses in a host specificity determinant. Science *357*, 80–83.

Ishii, H., Tanoue, J., Oshima, M., Chung, W.-H., Nishimura, K., Yamaguchi, J., Nemoto, F., So, K., Iwama, T., Yoshimatsu, H., et al. (2008). First application of PCR-Luminex system for molecular diagnosis of fungicide resistance and species identification of fungal pathogens. J. Gen. Plant Pathol. *74*, 409–416.

Islam, M.T., Croll, D., Gladieux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, Md.S., Gupta, D.R., Rahman, Md.M., Mahboob, M.G., et al. (2016). Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae. BMC Biol. *14*, 84.

Islam, M.T., Kim, K.-H., and Choi, and J. (2019). Wheat Blast in Bangladesh: The Current Situation and Future Impacts. Plant Pathol. J. *35*, 1–10.

James, T.Y. (2015). Why mushrooms have evolved to be so promiscuous: Insights from evolutionary and ecological patterns. Fungal Biol. Rev. *29*, 167–178.

Javan-Nikkhah, M., McDonald, B.A., Banke, S., and Hedjaroude, G.A. (2004). Genetic structure of Iranian *Pyricularia grisea* populations based on rep-PCR fingerprinting. Eur. J. Plant Pathol. *110*, 909–919.

Jiggins, null, and Mallet, null (2000). Bimodal hybrid zones and speciation. Trends Ecol. Evol. *15*, 250–255.

Jinks, J.L., and Grindle, M. (1963). The genetical basis oh heterokaryon incompatibility in *Aspergillus nidulans*. Heredity *18*, 407–411.

Johansson, H., Seppä, P., Helanterä, H., Trontti, K., and Sundström, L. (2018). Weak population structure in the ant *Formica fusca*. PeerJ *6*, e5024.

Johnson, J.A., Harrington, T.C., and Engelbrecht, C.J.B. (2005). Phylogeny and taxonomy of the North American clade of the *Ceratocystis fimbriata* complex. Mycologia *97*, 1067–1092.

Kato, H. (Kobe U. (Japan) F. of A., Yamamoto, M., Yamaguchi Ozaki, T., Kadouchi, H., Iwamoto, Y., Nakayashiki, H., Tosa, Y., Mayama, S., and Mori, N. (2000). Pathogenicity, mating ability and DNA restriction fragment length polymorphisms of *Pyricularia* populations isolated from Gramineae, Bambusideae and Zingiberaceae palnts. J. Gen. Plant Pathol. Jpn.

Khush, G.S. (2005). What it will take to feed 5.0 billion rice consumers in 2030. Plant Mol. Biol. *59*, 1–6.

Klaubauf, S., Tharreau, D., Fournier, E., Groenewald, J.Z., Crous, P.W., de Vries, R.P., and Lebrun, M.-H. (2014). Resolving the polyphyletic nature of *Pyricularia* (*Pyriculariaceae*). Stud. Mycol. *79*, 85–120.

Kohli, M.M., Mehta, Y.R., Guzman, E., De Viedma, L., and Cubilla, L.E. (2011). *Pyricularia* blast - a threat to wheat cultivation. Czech J. Genet. Plant Breed. *47*, S130–S134.

Koufopanou, V., Burt, A., Szaro, T., and Taylor, J.W. (2001). Gene Genealogies, Cryptic Species, and Molecular Evolution in the Human Pathogen *Coccidioides immitis* and Relatives (*Ascomycota, Onygenales*). Mol. Biol. Evol. *18*, 1246–1258.

Kroken, S., Glass, N.L., Taylor, J.W., Yoder, O.C., and Turgeon, B.G. (2003). Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc. Natl. Acad. Sci. U. S. A. *100*, 15670–15675.

Kumar, J., Nelson, R.J., and Zeigler, R.S. (1999). Population structure and dynamics of *Magnaporthe grisea* in the Indian Himalayas. Genetics *152*, 971–984.

Kuzdraliński, A., Kot, A., Szczerba, H., Nowak, M., and Muszyńska, M. (2017). A Review of Conventional PCR Assays for the Detection of Selected Phytopathogens of Wheat. J. Mol. Microbiol. Biotechnol. *27*, 175–189.

Lambert, D., and Spencer, H. (1996). Speciation and the Recognition Concept: Theory and Application. ,. Q. Rev. Biol. *71*, 274–275.

Le Gac, M., Hood, M.E., Fournier, E., and Giraud, T. (2007). Phylogenetic evidence of host-specific cryptic species in the anther smut fungus. Evol. Int. J. Org. Evol. *61*, 15–26.

Leducq, J.-B., Nielly-Thibault, L., Charron, G., Eberlein, C., Verta, J.-P., Samani, P., Sylvester, K., Hittinger, C.T., Bell, G., and Landry, C.R. (2016). Speciation driven by hybridization and chromosomal plasticity in a wild yeast. Nat. Microbiol. *1*, 15003.

Lee, S.C., Ni, M., Li, W., Shertz, C., and Heitman, J. (2010). The evolution of sex: a perspective from the fungal kingdom. Microbiol. Mol. Biol. Rev. MMBR *74*, 298–340.

Levy, M., Correa Victoria, F.J., Zeigler, R.S., Xu, S., and Hamer, J.E. (1993). Genetic diversity of the rice blast fungus in a disease nursery in Colombia. Phytopathology.

Liao, J., Huang, H., Meusnier, I., Adreit, H., Ducasse, A., Bonnot, F., Pan, L., He, X., Kroj, T., Fournier, E., et al. (2016). Pathogen effectors and plant immunity determine specialization of the blast fungus to rice subspecies. ELife *5*, e19377.

Lievens, B., and Thomma, B.P.H.J. (2005). Recent Developments in Pathogen Detection Arrays: Implications for Fungal Plant Pathogens and Use in Practice. Phytopathology *95*, 1374–1380.

M. L. Masters, J.C. (1987). The concept of species: Recognition versus Isolation. South Afr. J. Sci. *83*, 534.

Magdama, F., Monserrate-Maggi, L., Serrano, L., Sosa, D., Geiser, D.M., and Jiménez-Gasco, M. del M. (2019). Comparative analysis uncovers the limitations of current molecular detection methods for *Fusarium oxysporum f. sp. cubense* race 4 strains. PLOS ONE *14*, e0222727.

Mallet, J. (1995). A species definition for the modern synthesis. Trends Ecol. Evol. *10*, 294–299.

Mallet, J., Beltrán, M., Neukirchen, W., and Linares, M. (2007). Natural hybridization in *heliconiine* butterflies: the species boundary as a continuum. BMC Evol. Biol. *7*, 28.

Mancini, V., Murolo, S., and Romanazzi, G. (2016). Diagnostic methods for detecting fungal pathogens on vegetable seeds. Plant Pathol. *65*, 691–703.

Mayr, E. (1999). Systematics and the Origin of Species, from the Viewpoint of a Zoologist (Harvard University Press).

McManus, B.A., and Coleman, D.C. (2014). Molecular epidemiology, phylogeny and evolution of *Candida albicans*. Infect. Genet. Evol. *21*, 166–178.

McTaggart, A.R., van der Nest, M.A., Steenkamp, E.T., Roux, J., Slippers, B., Shuey, L.S., Wingfield, M.J., and Drenth, A. (2016). Fungal Genomics Challenges the Dogma of Name-Based Biosecurity. PLoS Pathog. *12*.

Merrill, R.M., Gompert, Z., Dembeck, L.M., Kronforst, M.R., McMillan, W.O., and Jiggins, C.D. (2011). Mate preference across the speciation continuum in a clade of mimetic butterflies. Evol. Int. J. Org. Evol. *65*, 1489–1500.

Michener, C.D. (1970). Diverse Approaches to Systematics.

Milazzo, J., Pordel, A., Ravel, S., and Tharreau, D. (2019). First Scientific Report of *Pyricularia oryzae* Causing Gray Leaf Spot Disease on Perennial Ryegrass (*Lolium perenne*) in France. Plant Dis. PDIS-09-18-1545.

Mishler, B.D. (1985). The Morphological, Developmental, and Phylogenetic Basis of Species Concepts in Bryophytes. The Bryologist *88*, 207–214.

Mumford, R., Boonham, N., Tomlinson, J., and Barker, I. (2006). Advances in molecular phytodiagnostics – new solutions for old problems. Eur. J. Plant Pathol. *116*, 1–19.

Murakishi, H.H. (1951). Purple seed stain of Soybean. Phytopathology *41*.

Murphy, H.A., Kuehne, H.A., Francis, C.A., and Sniegowski, P.D. (2006). Mate choice assays and mating propensity differences in natural yeast populations. Biol. Lett. *2*, 553–556.

Nelson, G., and Platnick, N. (1982). Review of Systematics and Biogeography: Cladistics and Vicariance. Syst. Zool. *31*, 206–208.

Nicolaisen, M., Justesen, A.F., Knorr, K., Wang, J., and Pinnschmidt, H.O. (2014). Fungal communities in wheat grain show significant co-existence patterns among species. Fungal Ecol. *11*, 145–153.

Nieuwenhuis, B.P.S., and James, T.Y. (2016). The frequency of sex in fungi. Philos. Trans. R. Soc. B Biol. Sci. *371*, 20150540.

Nieuwenhuis, B.P.S., Billiard, S., Vuilleumier, S., Petit, E., Hood, M.E., and Giraud, T. (2013). Evolution of uni- and bifactorial sexual compatibility systems in fungi. Heredity *111*, 445–455.

Nixon, K.C., and Wheeler, Q.D. (1990). An Amplification of the Phylogenetic Species Concept. Cladistics *6*, 211–223.

Nosil, P. (2007). Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. Am. Nat. *169*, 151–162.

Nosil, P. (2012). Ecological Speciation (OUP Oxford).

Onaga, G., Wydra, K., Koopmann, B., Sere, Y., and von Tiedemann, A. (2015). Population Structure, Pathogenicity, and Mating Type Distribution of *Magnaporthe oryzae* Isolates from East Africa. Phytopathology *105*, 1137–1145.

Orr-Weaver, T.L., and Szostak, J.W. (1985). Fungal recombination. Microbiol. Rev. *49*, 33–58.

Ou, S.H. (1980). Pathogen Variability and Host Resistance in Rice Blast Disease. Annu. Rev. Phytopathol. *18*, 167–187.

Ou, S.H. (1985). Rice Diseases (IRRI).

Pagliaccia, D., Urak, R.Z., Wong, F., Douhan, L.I., Greer, C.A., Vidalakis, G., and Douhan, G.W. (2018). Genetic Structure of the Rice Blast Pathogen (*Magnaporthe oryzae*) over a Decade in North Central California Rice Fields. Microb. Ecol. *75*, 310–317.

Park, S.Y., Milgroom, M.G., Han, S.S., Kang, S., and Lee, Y.H. (2003). Diversity of pathotypes and DNA fingerprint haplotypes in populations of Magnaporthe grisea in Korea over two decades. Phytopathology *93*, 1378–1385.

Park, S.-Y., Milgroorn, M.G., Han, S.S., Kang, S., and Lee, Y.-H. (2008). Genetic differentiation of Magnaporthe oryzae populations from scouting plots and commercial rice fields in Korea. Phytopathology *98*, 436–442.

Paterson, H.E.H. (1985). The Recognition Concept of Species. In Species and Speciation, E. Vrba, ed. (Transvaal Museum Monograph No. 4. Pretoria), p.

Perkins, D.D. (1988). Main features of vegetative incompatibility in Neurospora. Fungal Genet. Rep. *35*, 44.

Pieck, M.L., Ruck, A., Farman, M.L., Peterson, G.L., Stack, J.P., Valent, B., and Pedley, K.F. (2017). Genomics-Based Marker Discovery and Diagnostic Assay Development for Wheat Blast. Plant Dis. *101*, 103–109.

Pitt, J.I. (1979). The genus Penicillium and its teleomorphic states Eupenicillium and Talaromyces. Genus Penicillium Its Teleomorphic States Eupenicillium Talaromyces.

Qin, G.-F., Zhao, J., and Korhonen, K. (2007). A study on intersterility groups of Armillaria in China. Mycologia *99*, 430–441.

Randall-Schadel, B.L., Bailey, J.E., and Beute, M.K. (2001). Seed Transmission of *Cylindrocladium parasiticum* in Peanut. Plant Dis. *85*, 362–370.

Richards, T.A., Soanes, D.M., Foster, P.G., Leonard, G., Thornton, C.R., and Talbot, N.J. (2009). Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. Plant Cell *21*, 1897–1911.

Richards, T.A., Soanes, D.M., Jones, M.D.M., Vasieva, O., Leonard, G., Paszkiewicz, K., Foster, P.G., Hall, N., and Talbot, N.J. (2011). Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. Proc. Natl. Acad. Sci. *108*, 15258–15263.

Rieux, A., and Balloux, F. (2016). Inferences from tip-calibrated phylogenies: a review and a practical guide. Mol. Ecol. *25*, 1911–1924.

Rosen, D.E. (1979). Fishes from the uplands and intermontane basins of Guatemala : revisionary studies and comparative geography. Bulletin of the AMNH ; v. 162, article 5.

Roumen, E., Levy, M., and Nottéghem, J.-L. (1997). Characterisation of the European pathogen population of *Magnaporthe grisea* by DNA fingerprinting and pathotype analysis. Eur. J. Plant Pathol.

Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., and Bierne, N. (2016). Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. PLOS Biol. *14*, e2000234.

Rundell, R.J., and Price, T.D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. Trends Ecol. Evol. *24*, 394–399.

Rundle, H.D., and Nosil, P. (2005). Ecological speciation. Ecol. Lett. *8*, 336–352.

Rymer, P.D., Manning, J.C., Goldblatt, P., Powell, M.P., and Savolainen, V. (2010). Evidence of recent and continuous speciation in a biodiversity hotspot: a population genetic approach in southern African *gladioli* (*Gladiolus; Iridaceae*). Mol. Ecol. *19*, 4765–4782.

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigné, V., Bazin, E., Nottéghem, J.-L., Fournier, E., et al. (2012). Sex at the origin: an Asian population of the rice blast fungus *Magnaporthe oryzae* reproduces sexually. Mol. Ecol. *21*, 1330–1344.

Saleh, D., Milazzo, J., Adreit, H., Fournier, E., and Tharreau, D. (2014). South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. New Phytol. *201*, 1440–1456.

Samils, B., Stepien, V., Lagercrantz, U., Lascoux, M. and Gullberg, U. Genetic Diversity in Relation to Sexual and Asexual Reproduction in Populations of Melampsora Larici-Epitea. European Journal of Plant Pathology *107,* 871-81.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. Nat. Ecol. Evol. *3*, 430–439.

Schena, L., Nigro, F., Ippolito, A., and Gallitelli, D. (2004). Real-time quantitative PCR: a new technology to detect and study phytopathogenic and antagonistic fungi. Eur. J. Plant Pathol. *110*, 893–908.

Schluter, D. (2001). Ecology and the origin of species. Trends Ecol. Evol. *16*, 372–380.

Schluter, D. (2009). Evidence for ecological speciation and its alternative. Science *323*, 737–741.

Shoji, J.-Y., Charlton, N.D., Yi, M., Young, C.A., and Craven, K.D. (2015). Vegetative hyphal fusion and subsequent nuclear behavior in *Epichloë* grass endophytes. PloS One *10*, e0121875.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. Science *236*, 787–792.

Sneath, P.H.A., and Sokal, R.R. (1973). Numerical taxonomy: the principles and practice of numerical classification (San Francisco: W. H. Freeman).

Soanes, D., and Richards, T.A. (2014). Horizontal Gene Transfer in Eukaryotic Plant Pathogens. Annu. Rev. Phytopathol. *52*, 583–614.

Steenkamp, E.T., Wingfield, M.J., McTaggart, A.R., and Wingfield, B.D. (2018). Fungal species and their boundaries matter – Definitions, mechanisms and practical implications. Fungal Biol. Rev. *32*, 104–116.

Stukenbrock, E.H., Christiansen, F.B., Hansen, T.T., Dutheil, J.Y., and Schierup, M.H. (2012). Fusion of two divergent fungal individuals led to the recent emergence of a unique widespread pathogen species. Proc. Natl. Acad. Sci. U. S. A. *109*, 10954–10959.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S., and Fisher, M.C. (2000). Phylogenetic species recognition and species concepts in fungi. Fungal Genet. Biol. FG B *31*, 21–32.

Taylor, J.W., Branco, S., Gao, C., Hann-Soden, C., Montoya, L., Sylvain, I., and Gladieux, P. (2017). Sources of Fungal Genetic Variation and Associating It with Phenotypic Diversity. Microbiol. Spectr. *5*.

Taylor John W, Turner Elizabeth, Townsend Jeffrey P, Dettman Jeremy R, and Jacobson David (2006). Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. Philos. Trans. R. Soc. B Biol. Sci. *361*, 1947–1963.

Templeton, A.R. (2008). The reality and importance of founder speciation in evolution. BioEssays News Rev. Mol. Cell. Dev. Biol. *30*, 470–479.

Tharreau, D., Fudal, I., Andriantsimialona, D., Santoso, Utami, D., Fournier, E., Lebrun, M.-H., and Notteghem, J.-L. (2009). World Population Structure and Migration of the Rice Blast Fungus, *Magnaporthe oryzae* (Dordrecht: Springer).

Thierry, M., Gladieux, P., Fournier, E., Tharreau, D., and Ioos, R. (2019). A genomic approach to develop a new qPCR test enabling detection of the *Pyricularia oryzae* lineage causing wheat blast. Plant Dis.

Thorpe, R.S., Surget-Groba, Y., and Johansson, H. (2010). Genetic tests for ecological and allopatric speciation in anoles on an island archipelago. PLoS Genet. *6*, e1000929.

Thuan, N.T.N., Bigirimana, J., Roumen, E., Van Der Straeten, D., and Höfte, M. (2006). Molecular and Pathotype Analysis of the Rice Blast Fungus in North Vietnam. Eur. J. Plant Pathol. *114*, 381–396.

Tosa, Y., Inoue, Y., Trinh, T.P.V., and Chuma, I. (2016). Genetic and molecular analyses of the incompatibility between Lolium isolates of Pyricularia oryzae and wheat. Physiol. Mol. Plant Pathol. *95*, 84–86.

Treindl, A.D., and Leuchtmann, A. (2019). Assortative mating in sympatric ascomycete fungi revealed by experimental fertilizations. Fungal Biol. *123*, 676–686.

Valen, L.V. (1976). Ecological Species, Multispecies, and Oaks. Taxon *25*, 233–239.

Valent, B., Farman, M., Tosa, Y., Begerow, D., Fournier, E., Gladieux, P., Islam, M.T., Kamoun, S., Kemler, M., Kohn, L.M., et al. (2019). Pyricularia graminis-tritici is not the correct species name for the wheat blast fungus: response to Ceresini et al. (MPP 20:2). Mol. Plant Pathol. *20*, 173–179.

Van Putten, W.F., Elzinga, J.A., and Biere, A. (2007). Host Fidelity of the Pollinator Guilds of *Silene dioica* and *Silene latifolia*: Possible Consequences for Sympatric Host Race Differentiation of a Vectored Plant Disease. Int. J. Plant Sci. *168*, 421–434.

Viedma, L.Q., and Morel, W. (2002). Añublo o Piricularia del Trigo. Díptico. MAG/DIA/CRIA. Programa de Investi-gación de Trigo, CRIA, Capitán Miranda, Itapúa.

Wang, Y., White, M.M., Kvist, S., and Moncalvo, J.-M. (2016). Genome-Wide Survey of Gut Fungi (*Harpellales*) Reveals the First Horizontally Transferred Ubiquitin Gene from a Mosquito Host. Mol. Biol. Evol. *33*, 2544–2554.

Wang, X., Jia, Y., Wamishe, Y., Jia M.H. and Valent B. (2017). Dynamic Changes in the Rice Blast Population in the United States Over Six Decades. Mol. Plant-Microbe Interact. *30*, 803-12.

Ward, E., Foster, S.J., Fraaije, B.A., and McCartney, H.A. (2004). Plant pathogen diagnostics: immunological and nucleic acid-based approaches. Ann. Appl. Biol.

Yasuhara-Bell, J., Pedley, K.F., Farman, M., Valent, B., and Stack, J.P. (2018). Specific detection of the wheat blast pathogen (*Magnaporthe oryzae* Triticum) by loop-mediated isothermal amplification. Plant Dis. 102, 2550-2559.

Zeigler, R.S. (1998). Recombination in *Magnaporthe grisea*. Annu. Rev. Phytopathol. *36*, 249–275.

Zhong, Z., Chen, M., Lin, L., Han, Y., Bao, J., Tang, W., Lin, L., Lin, Y., Somai, R., Lu, L., et al. (2018). Population genomic analysis of the rice blast fungus reveals specific events associated with expansion of three main clades. ISME J.

Diagnostic and inference of the evolutionary history of endemic and pandemic lineages of *Pyricularia oryzae* causing blast of rice, wheat and other Poaceae.

Abstract

Studying the genetic structure of populations within a species using genomic data provides insights into their evolution, the spread of individuals, the mode of reproduction or the existence of gene flow between populations. Regarding pathogenic species, understanding the genetic structure can be key for the management of epidemics or the prediction of their evolution. The fungal species *Pyricularia oryzae* infects many Poaceae species and is responsible for more than 4% yield loss on global rice production. In 1985, a host jump led to the emergence of wheat blast. Epidemics on this cereal were limited to South America but emerged very recently in Asia. The study of 81 genomes of *P. oryzae* collected from different species of Poaceae confirmed host-specific lineages. Differentiated genetic groups have been identified among the Oryza lineage (composed of rice pathogen isolates). The factors involved in this differentiation, however, are not known.

This work is focused on the population structure within *P. oryzae* and had two main objectives: i) to identify alleles specific to the Triticum lineage causing wheat blast epidemics in order to develop diagnostic tools to prevent the introduction of the pathogen in disease-free areas; ii) Increase knowledge about extant lineages within the Oryza lineage and identify the factors causing the genetic structure.

The long-range spread of wheat blast is attributed to the transport of contaminated biological material. The development of reliable pathogen detection methods is therefore crucial to avoid the importation of the pathogen into new areas. The previously established structure in *P. oryzae* highlighted the Triticum lineage as responsible for wheat blast epidemics. A genome comparison approach developed during this thesis identified multiple alleles highly specific of the Triticum lineage. These alleles were subsequently targeted for the development of molecular detection tests using several amplification techniques. Conventional PCR, real-time PCR and LAMP (Isothermal Amplification) assays, targeting several genomic regions, were developed and improved the detection of the pathogen. The application of these tests to the detection from infected seeds was shown.

To clarify the population structure within the rice-infecting Oryza lineage and to improve our knowledge of its evolutionary history, a dataset of 5000 SNPs characterized for 900 isolates covering the geographic range of the pathogen was analysed. Our analysis showed four major lineages, including two pandemic clonal lineages, a predominantly South Asian clonal lineage, and a recombinant lineage common in Southeast Asia. The recombinant lineage also had a geographical sub-structure. The barriers to gene flow allowing the maintenance of the genetic structure have been investigated. A strong correlation between lineage geographical range and climatic data was determined. Differential adaptations of these lineages to environmental factors were then evaluated and we identified differences in host spectra between lineages. Strong reproductive barriers, as well as a loss of fertility in some lineages, were also shown to limit the gene flow between lineages.

Key words: Genetic structure, diagnostic, blast, Poaceae, propagation, genomics

Diagnostic et inférence de l'histoire évolutive des lignées endémiques et pandémiques de *Pyricularia oryzae* causant la pyriculariose du riz, du blé, et d'autres Poacées sauvages

Résumé

L'étude de la structure génétique des populations d'une espèce à partir de données génomiques permet de comprendre son évolution, l'histoire de la propagation des individus, ses modes de reproduction ou les échanges de matériel génétique entre populations. Dans le cas d'espèces de microorganismes pathogènes, caractériser la structure et en comprendre les causes peut apporter des connaissances cruciales pour la gestion des épidémies ou la prédiction de leur évolution. Le champignon ascomycète *Pyricularia oryzae* infecte de nombreuses espèces de poacées et est responsable à lui seul de plus de 4% de perte de rendement sur la production mondiale de riz. En 1985, un saut d'hôte a entrainé l'émergence de la pyriculariose du blé. Les épidémies sur cette céréale restaient circonscrites à l'Amérique du sud mais ont émergé très récemment en Asie. L'étude de 81 génomes de *P. oryzae* prélevés de différentes espèces de poacées a confirmé l'existence de lignées hôte-spécifiques. Parmi celles-ci, la lignée Oryza (pathogène du riz) au sein de laquelle des groupes génétiques différenciés ont également été mis en évidence. Cependant, les facteurs impliqués dans cette différenciation ne sont pas connus.

Ce travail de thèse porte sur l'étude de la structure des populations au sein de *P. oryzae* et avait deux objectifs principaux : i) mettre en évidence des polymorphismes spécifiques de la lignée Triticum responsable des épidémies de pyriculariose du blé afin de développer des outils de diagnostic essentiels pour éviter l'introduction de l'agent pathogène dans des régions encore indemnes de cette maladie ; ii) Affiner nos connaissances sur les lignées existantes au sein de la lignée Oryza et déterminer les facteurs responsables de cette structure génétique.

La propagation sur de longues distances de la pyriculariose du blé est rendue possible par le transport de matériel biologique contaminé, comme les semences. La mise au point de méthodes de détection fiables de l'agent pathogène est donc cruciale pour éviter son importation dans de nouvelles régions. La structure en lignée hôte spécifique précédemment établie a mis en évidence la lignée Triticum regroupant les isolats responsables des épidémies de pyriculariose du blé. Une approche de comparaison de génomes développée au cours de cette thèse a permis d'identifier de multiples allèles spécifiques de la lignée Triticum. Ces allèles ont été ciblés pour le développement de tests moléculaires de détection utilisant plusieurs techniques d'amplification. Des tests de PCR conventionnelle, de PCR en temps-réel et LAMP (amplification isothermale) ciblant plusieurs régions génomiques ont été développés et améliorent la détection de l'agent pathogène. L'application de ces tests à la détection sur des semences infectées a également été vérifiée.

Afin de préciser la structure des populations au sein de la lignée infectant le riz (lignée Oryza) et d'améliorer nos connaissances sur son histoire évolutive un jeu de données de 5000 SNPs caractérisés chez 900 isolats couvrant l'aire de répartition de l'agent pathogène a été analysé. Nos analyses mettent en évidence quatre lignées majeures, dont deux lignées clonales pandémiques, une lignée clonale essentiellement Sud-Asiatique, et une lignée recombinante fréquente en Asie du Sud-Est. La lignée recombinante présente de plus une sous-structure géographique. Les barrières aux flux de gènes permettant le maintien de cette structure génétique ont été recherchées. Une forte corrélation entre aire de répartition des lignées et données climatiques a été mise en évidence. Des adaptations différentielles de ces lignées à des facteurs environnementaux (température et plante hôte) ont ensuite été évaluées identifiant des différences de largeur de spectres d'hôte entre lignées. De fortes barrières à la reproduction ont également été révélées, ainsi qu'une perte de fertilité dans certaines lignées, limitant le flux de gènes entre ces lignées.

Mots clefs : Structure génétique, diagnostic, pyriculariose, poacées, propagation, génomique